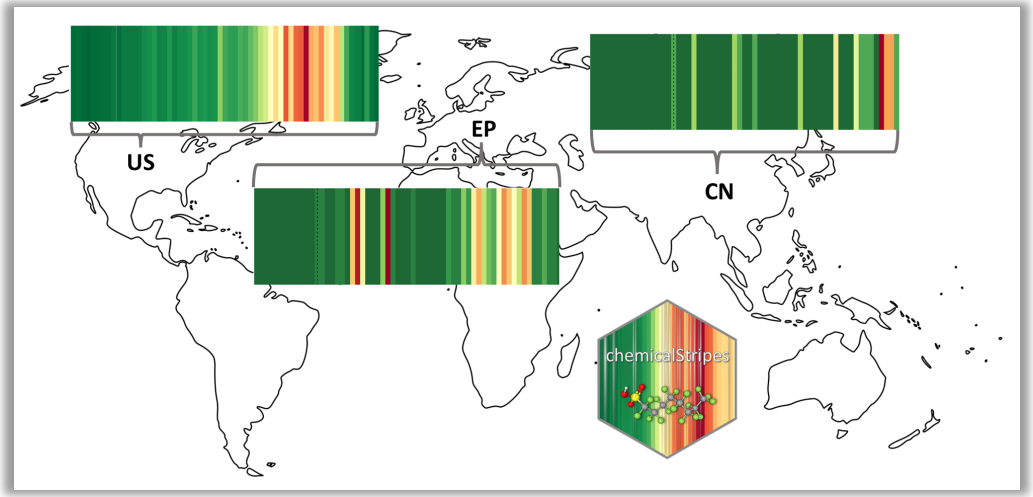
Luxembourg National Research Fund

Congressi Stefano Franscini

UNIVERSITÉ DU LUXEMBOURG

Funded by the European Union

Tracking Chemical Exposures in Time and Space



Prof. Dr. Emma L. Schymanski

(plus many, many colleagues and collaborators!)

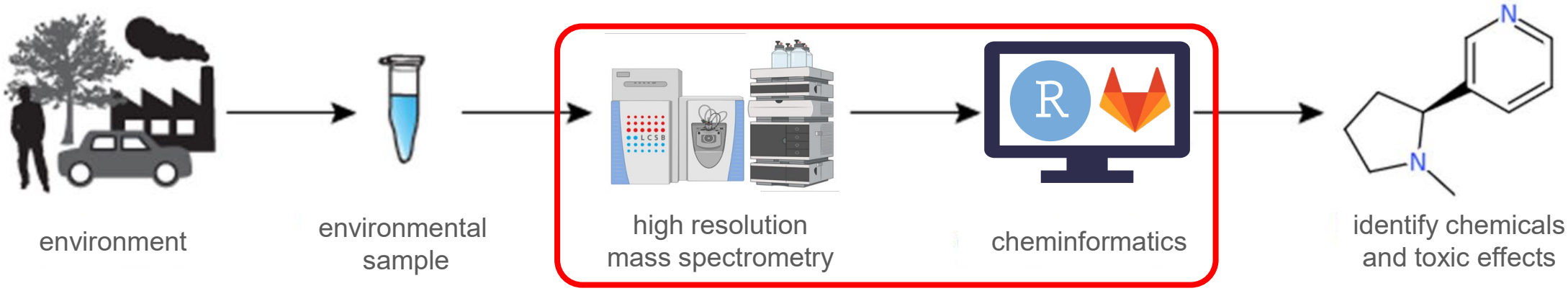
Environmental Cheminformatics Group,
Luxembourg Centre for Systems Biomedicine, University of Luxembourg
emma.schymanski@uni.lu / @ESchymanski / @schymane@mstdn.social
https://wwwen.uni.lu/lcsb/research/environmental_cheminformatics/

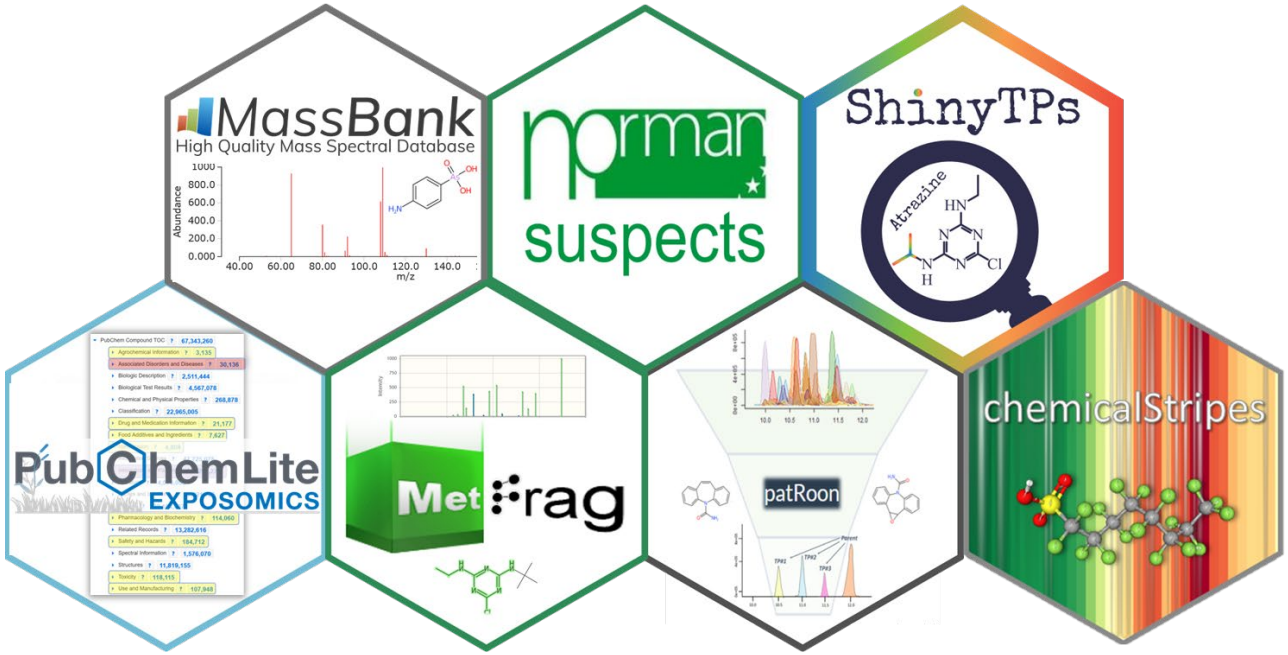
Keynote Speaker: Spatial and temporal statistical modeling in molecular biology, 8-13 September 2024, Ascona, CH

Slides available at DOI: 10.5281/zenodo.13755630

Today's slides zenodo

Environmental Cheminformatics & Non-target HR-MS

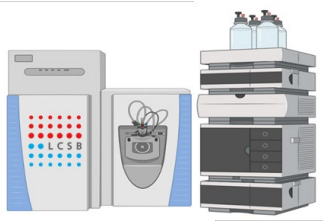




Environmental Cheminformatics & Non-target HR-MS

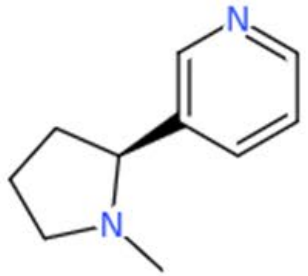
environment

environmental sample

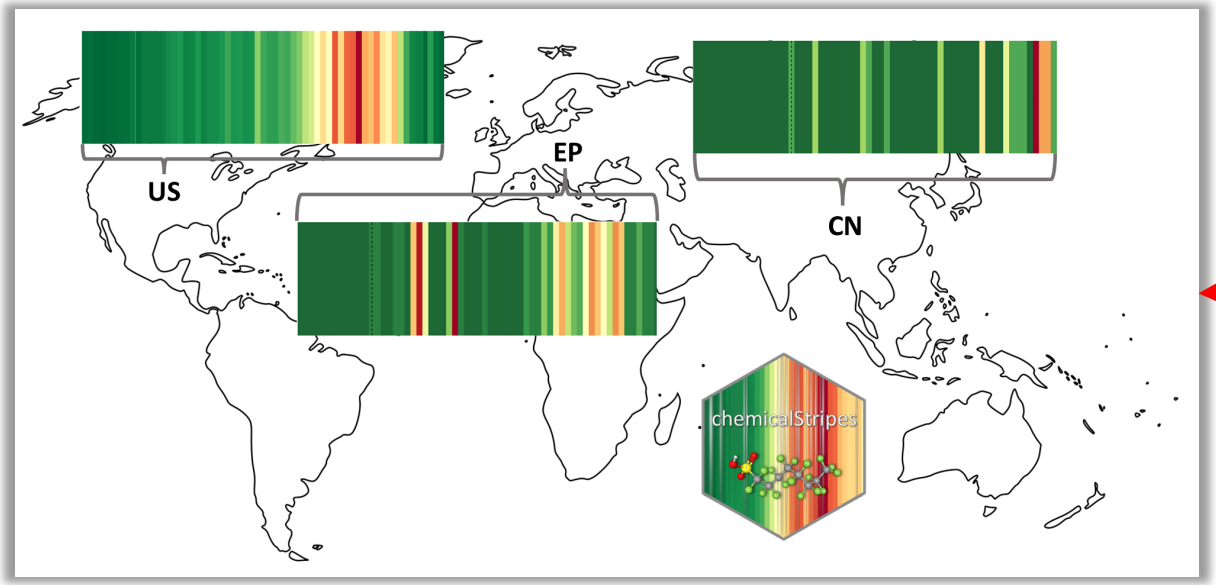


high resolution mass spectrometry

cheminformatics

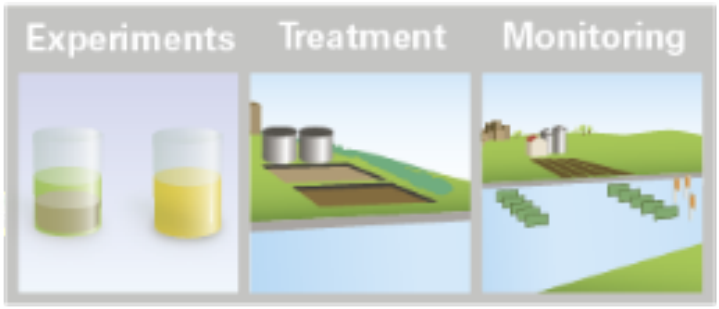


identify chemicals and toxic effects



…in space and time

Non-target High Resolution Mass Spectrometry (NT-HRMS)

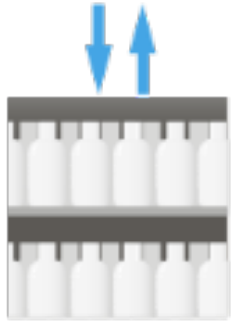


Commercial / Open Software

Databases

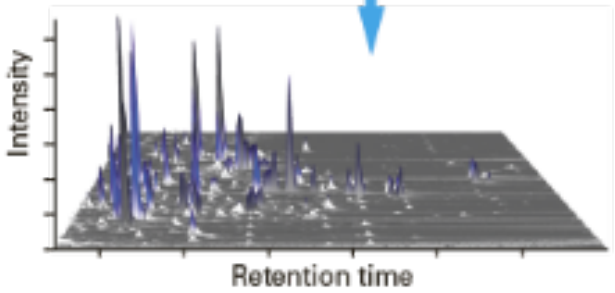
In Silico Prediction

Sampling → Analysis → Data Pre-Processing → Prioritization → Identification

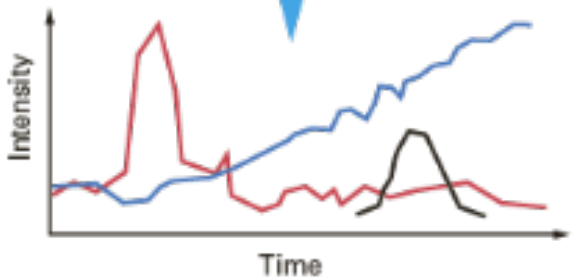


SAMPLE ARCHIVE

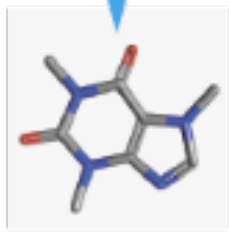
DIGITAL ARCHIVE



POLLUTION OVERVIEW



PATTERN

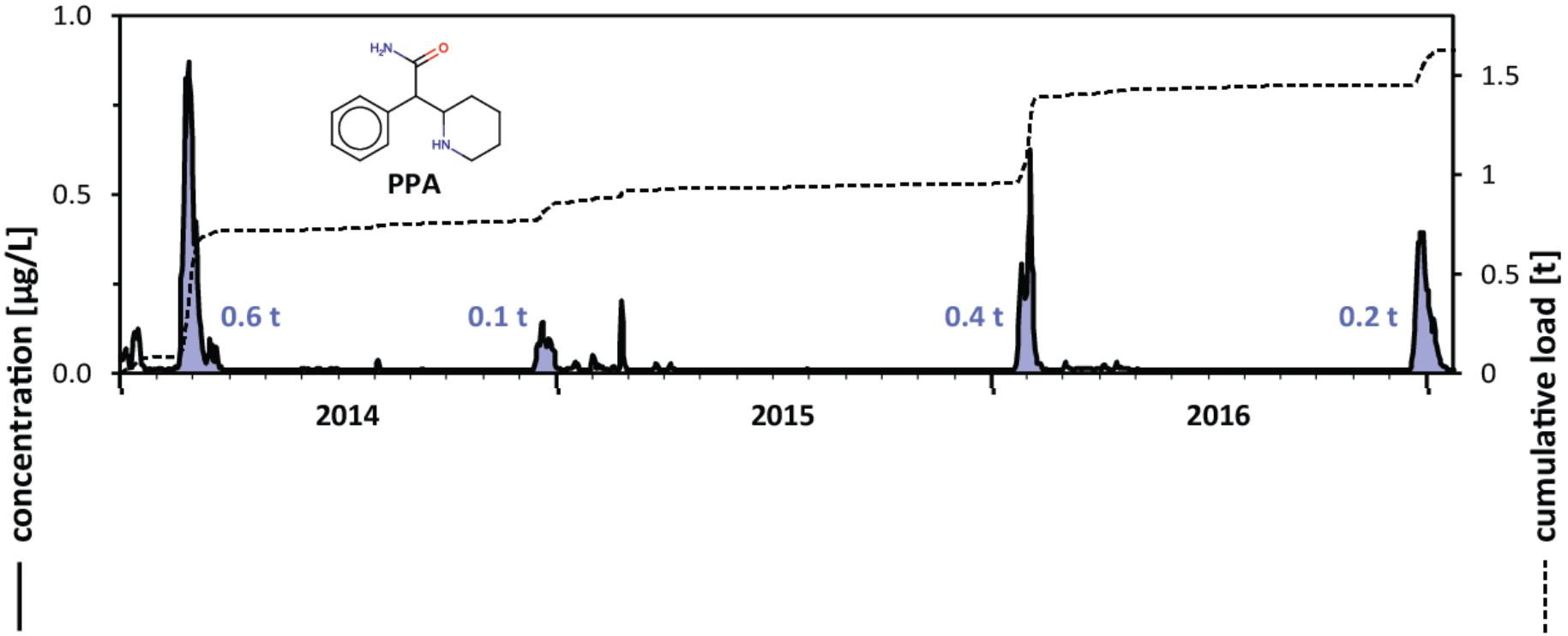


STRUCTURES

Modified from: Hollender, Schymanski *et al*. 2017, ES&T Feature, 51:20, 11505-11512. DOI: 10.1021/acs.est.7b02184

Examples: Real Time Monitoring of the Rhine River

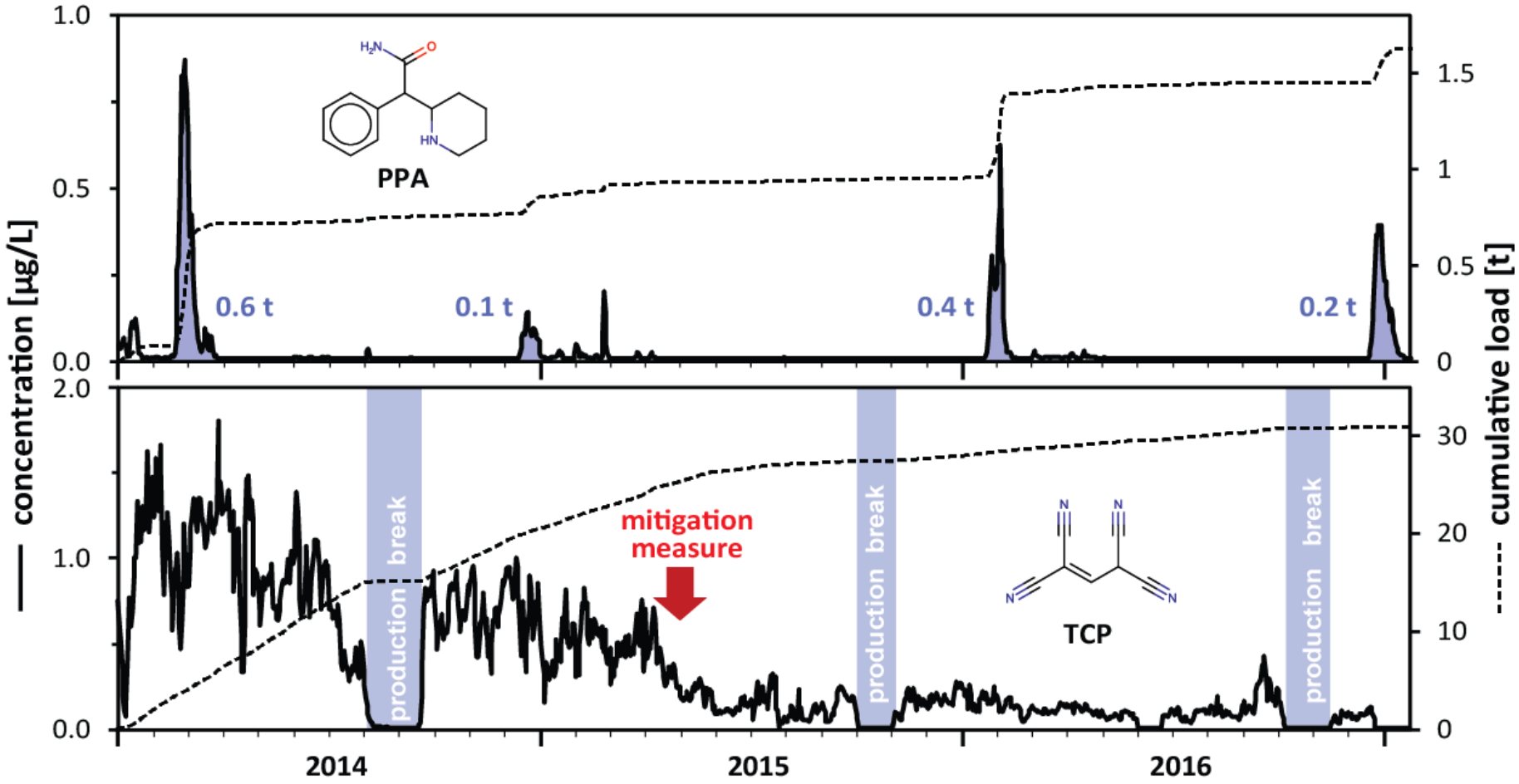
Previously unknown chemicals detected due to “stand-out” patterns



Hollender, Schymanski, Singer & Ferguson (2017) ES&T Feature, 51:20, 11505-11512. DOI: 10.1021/acs.est.7b02184

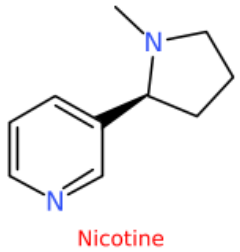
Examples: Real Time Monitoring of the Rhine River


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

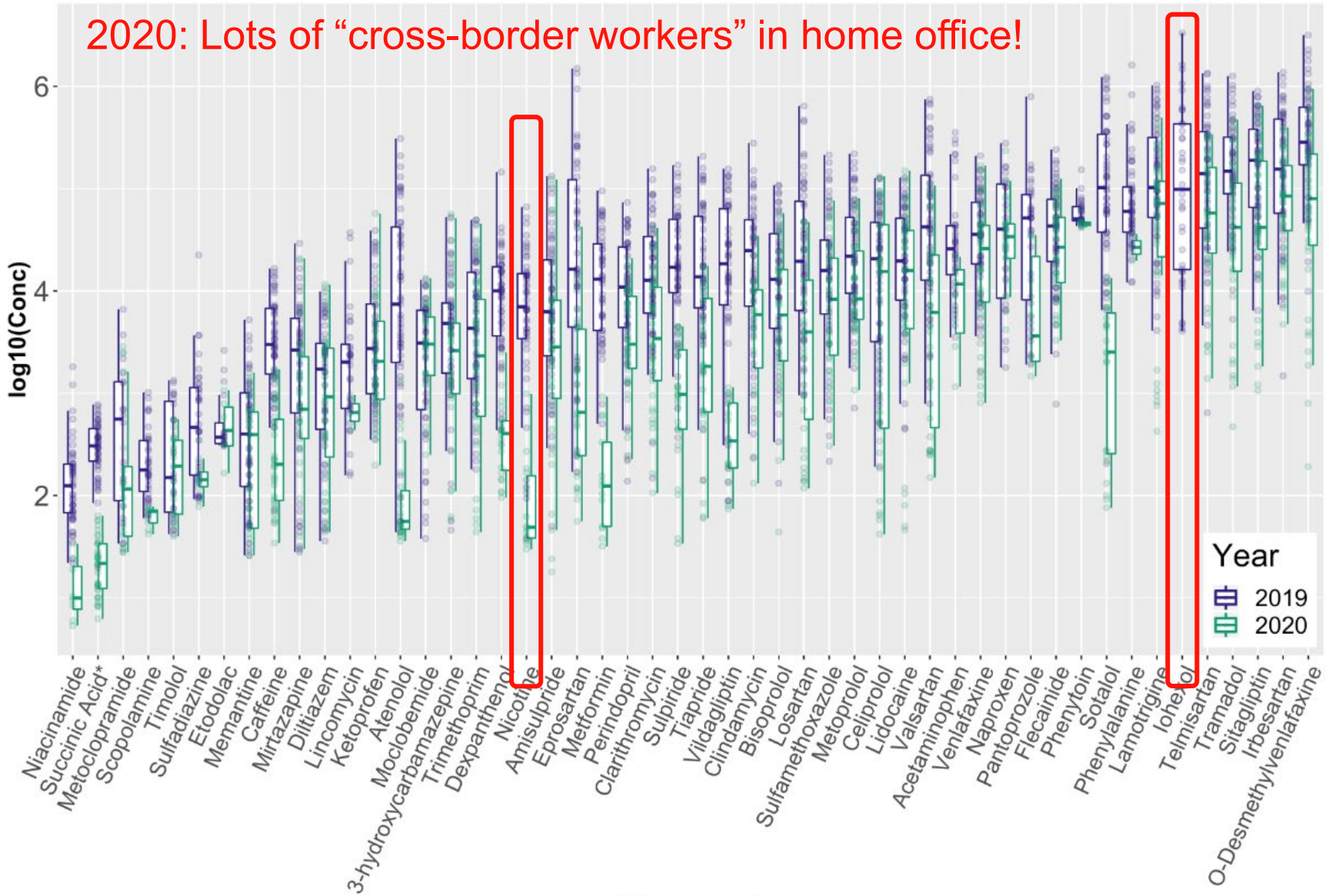


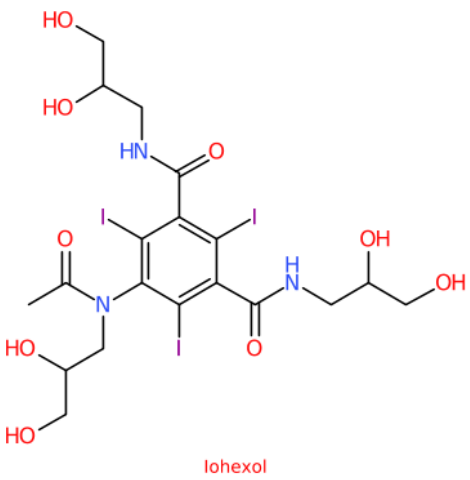
Hollender, Schymanski, Singer & Ferguson (2017) ES&T Feature, 51:20, 11505-11512. DOI: 10.1021/acs.est.7b02184

Examples: LuxPharma – 2019 versus 2020 & COVID?



Nicotine

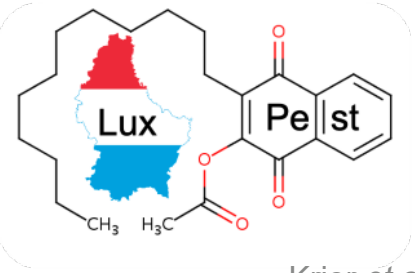




Iohexol

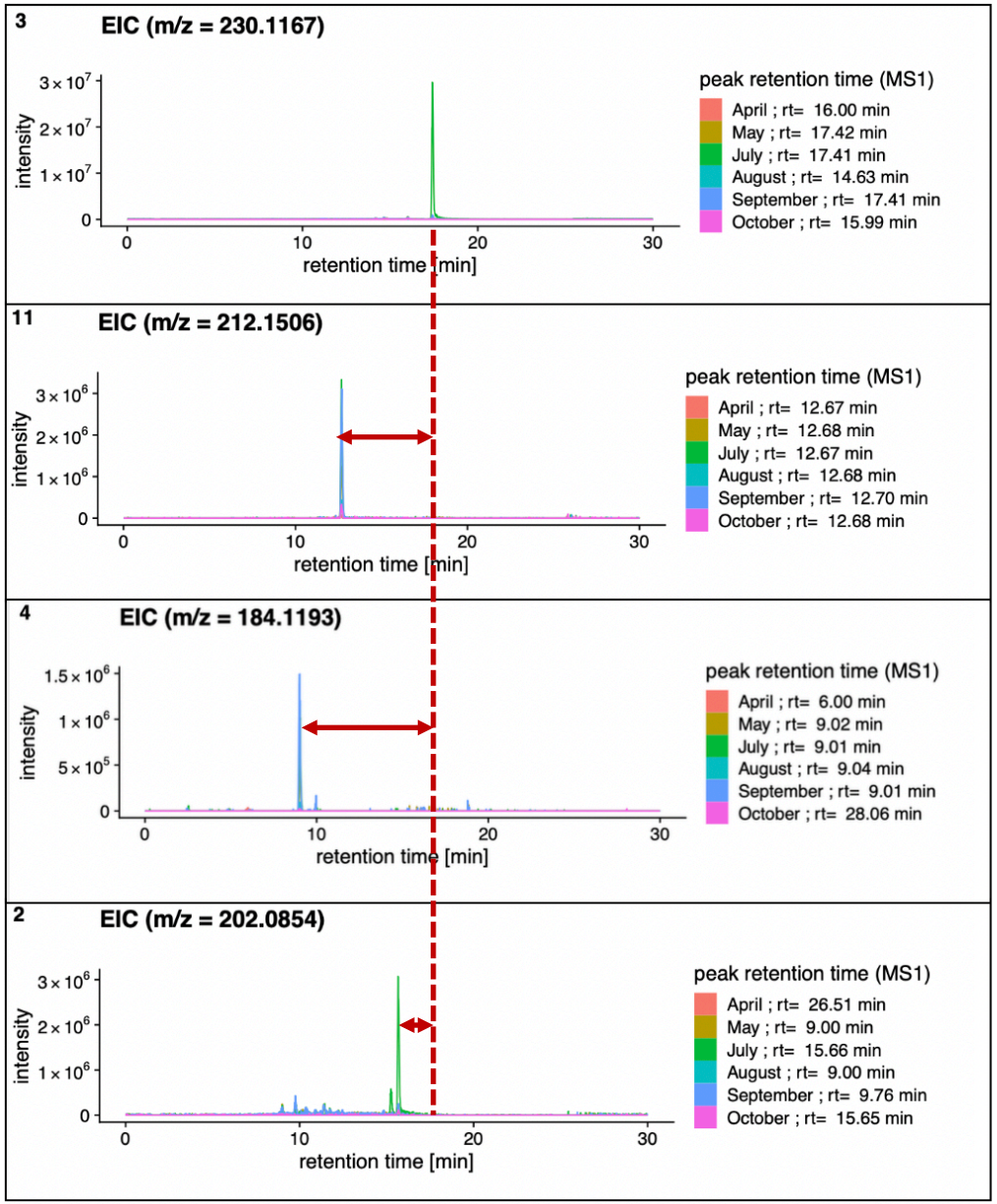
Singh *et al* (2021).
DOI: [10.1021/acsenvironau.1c00008](https://doi.org/10.1021/acsenvironau.1c00008).

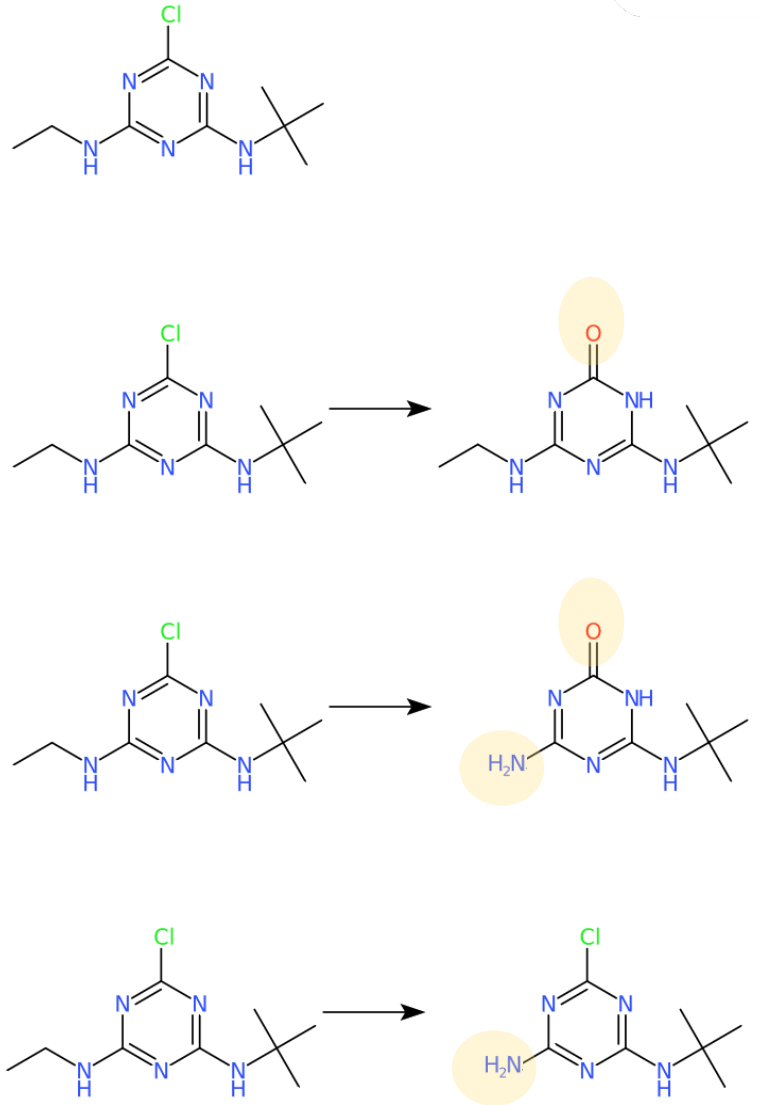
Examples: LuxPest and Transformation Products



Krier *et al* (2022). DOI: 10.1016/j.envint.2021.106885

THE GOVERNMENT
OF THE GRAND DUCHY OF LUXEMBOURG
Ministry of the Environment, Climate
and Sustainable Development





NORMAN DSFP – Digital Sample Freezing Platform

www.norman-data.eu

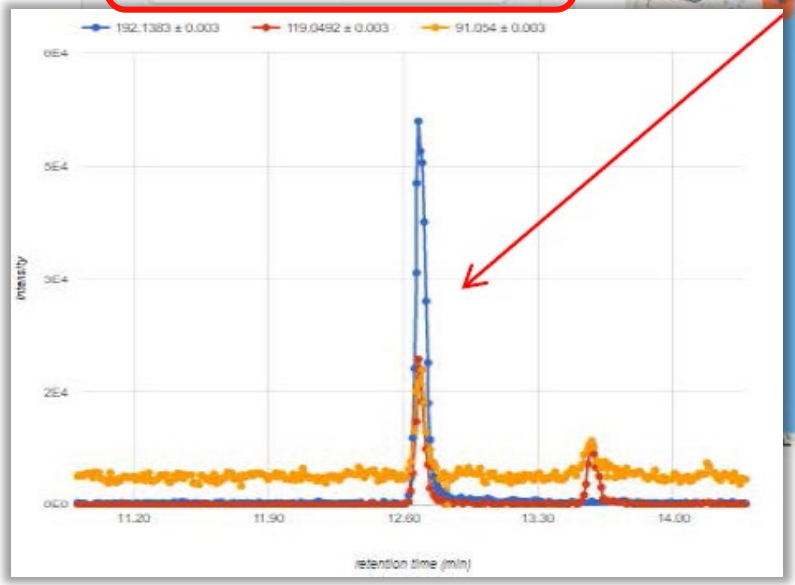
NORMAN Digital Sample Freezing Platform | Main Page | Batch mode

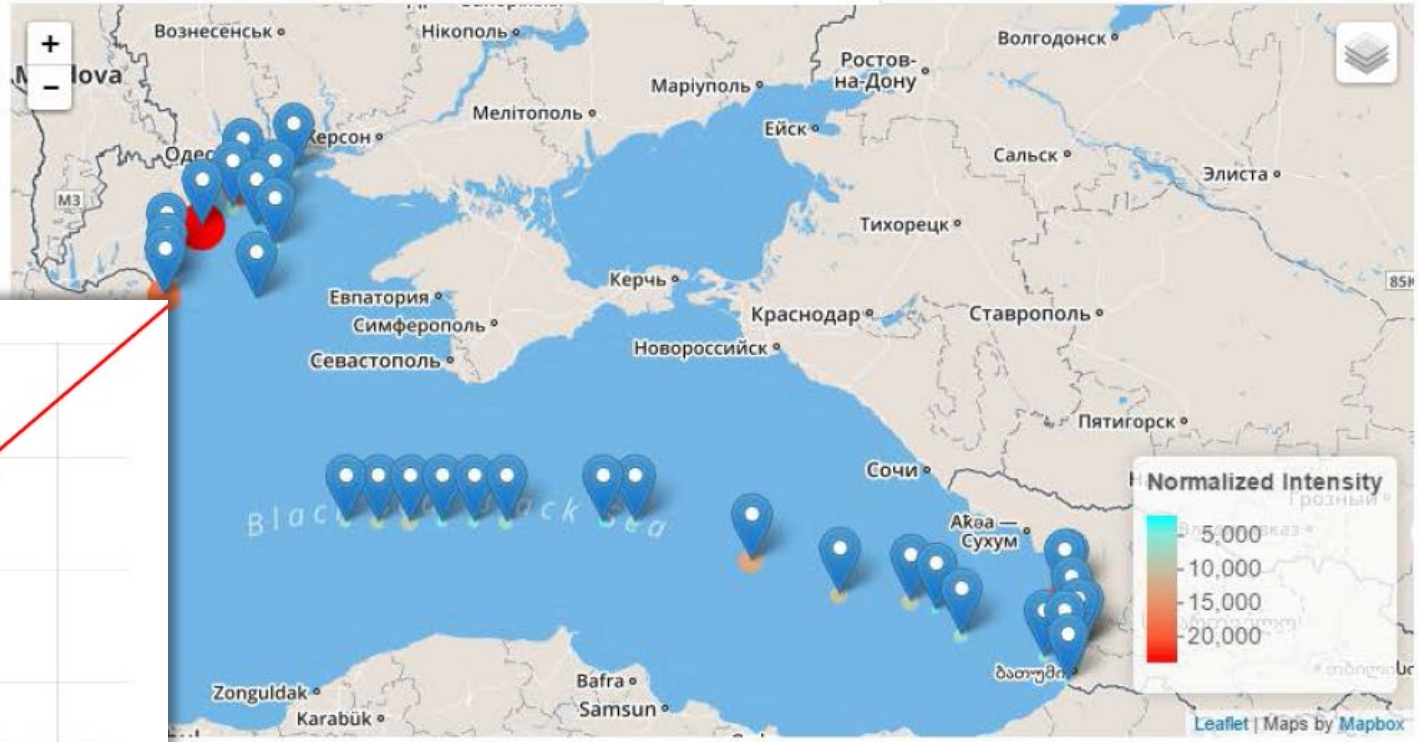
Choose Emerging Substance or input mass of interest and experimental RTI

Substance name or CAS or StdInChIKey

DEET [134-62-3] [MMOXZBCLCQITDF-UHFFFAOYSA-N]

Contributed Samples | Results | Chromatograms | Interactive Map | Help

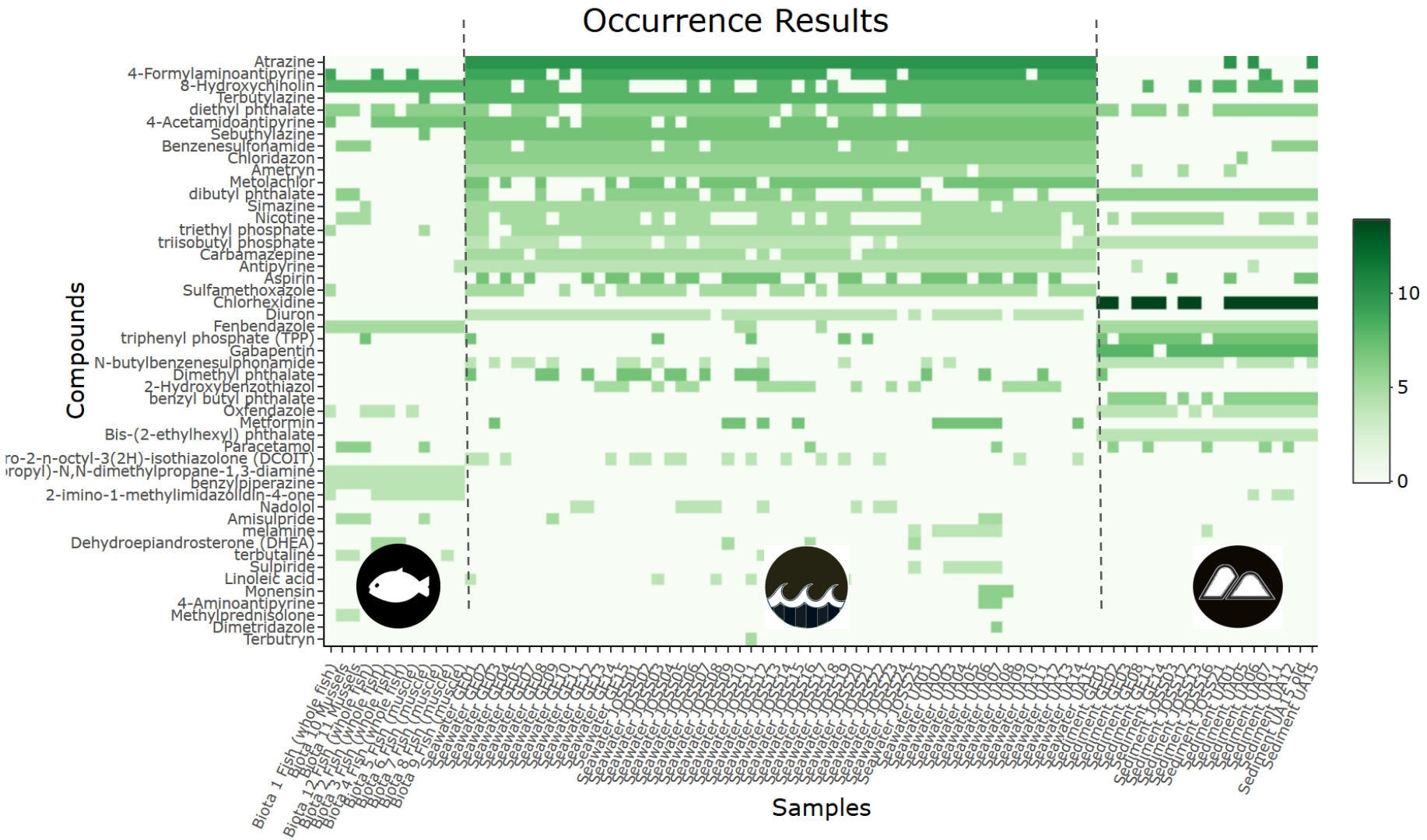




http://norman-data.eu/ and Alygizakis *et al*, 2019, TrAC, DOI: 10.1016/j.trac.2019.04.008

Examples: Distribution of Chemicals in Various Matrices

Retrospective screening of REACH chemicals in Black Sea samples



Alygizakis NA *et al* (2019) DOI: [10.1016/j.trac.2019.04.008](https://doi.org/10.1016/j.trac.2019.04.008)

Interactive heat map at http://norman-data.eu/NORMAN-REACH

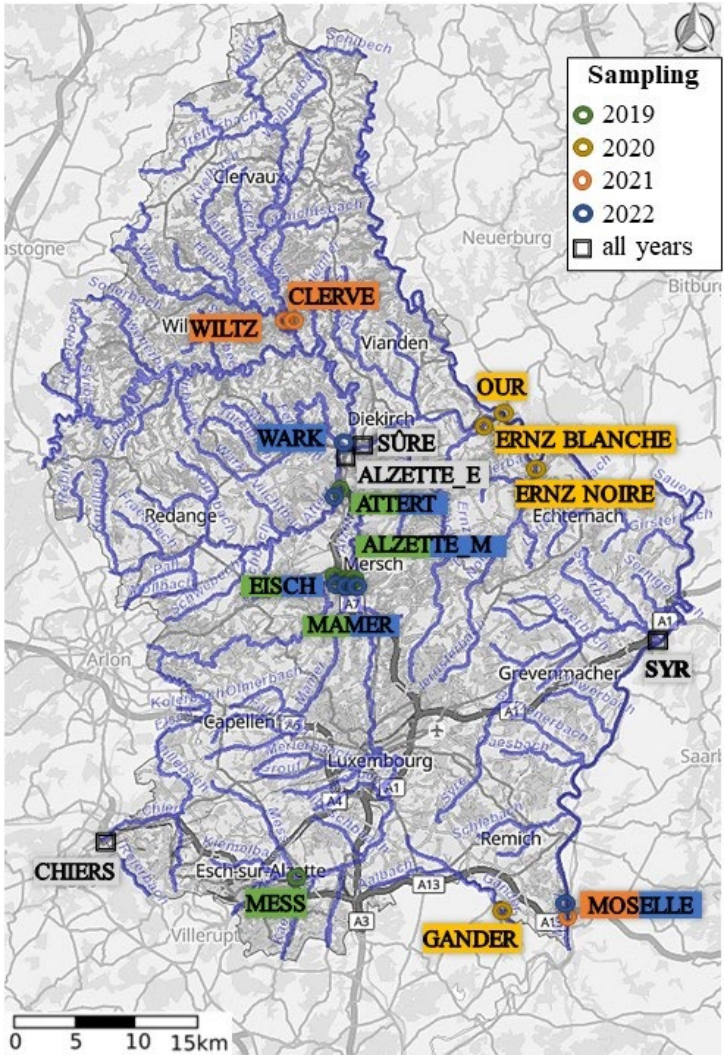
Latest NTS Results from Luxembourg

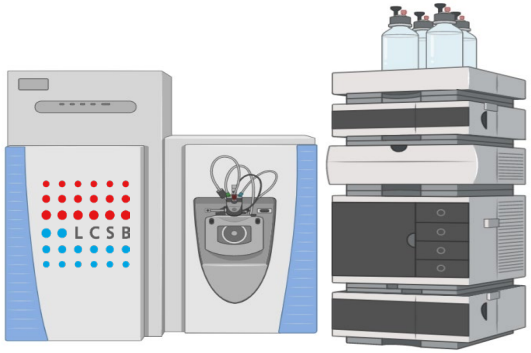
"No river in good condition" (2022)

271 surface water samples (2019-2022)

LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG
Ministère de l'Environnement, du Climat et du Développement durable

Administration de la gestion de l'eau





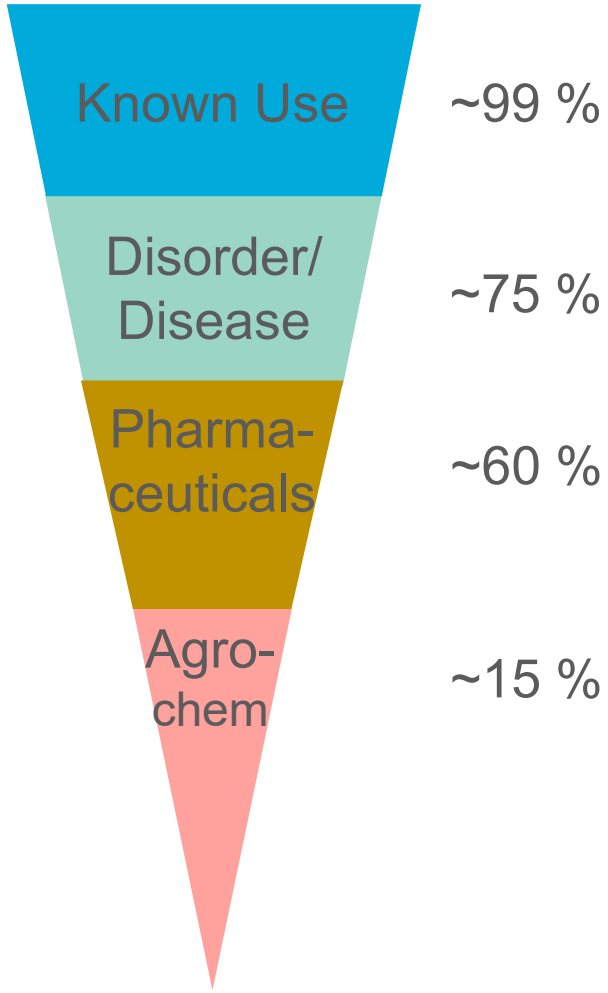
Target Monitoring

Complementary NTS

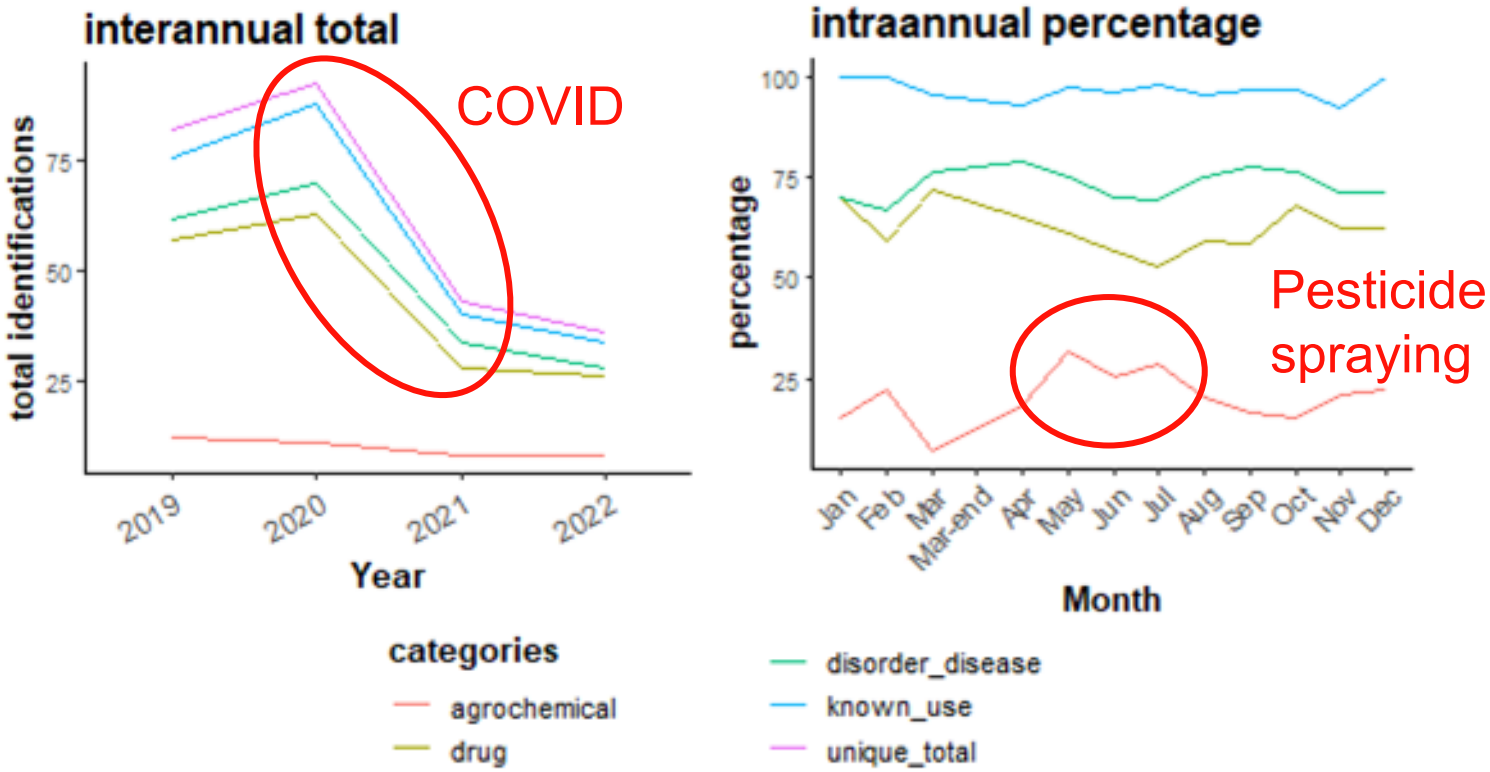
Slide: D. Aurich, PhD Defense. Ref: Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

Temporal Variations in Luxembourg

Classification

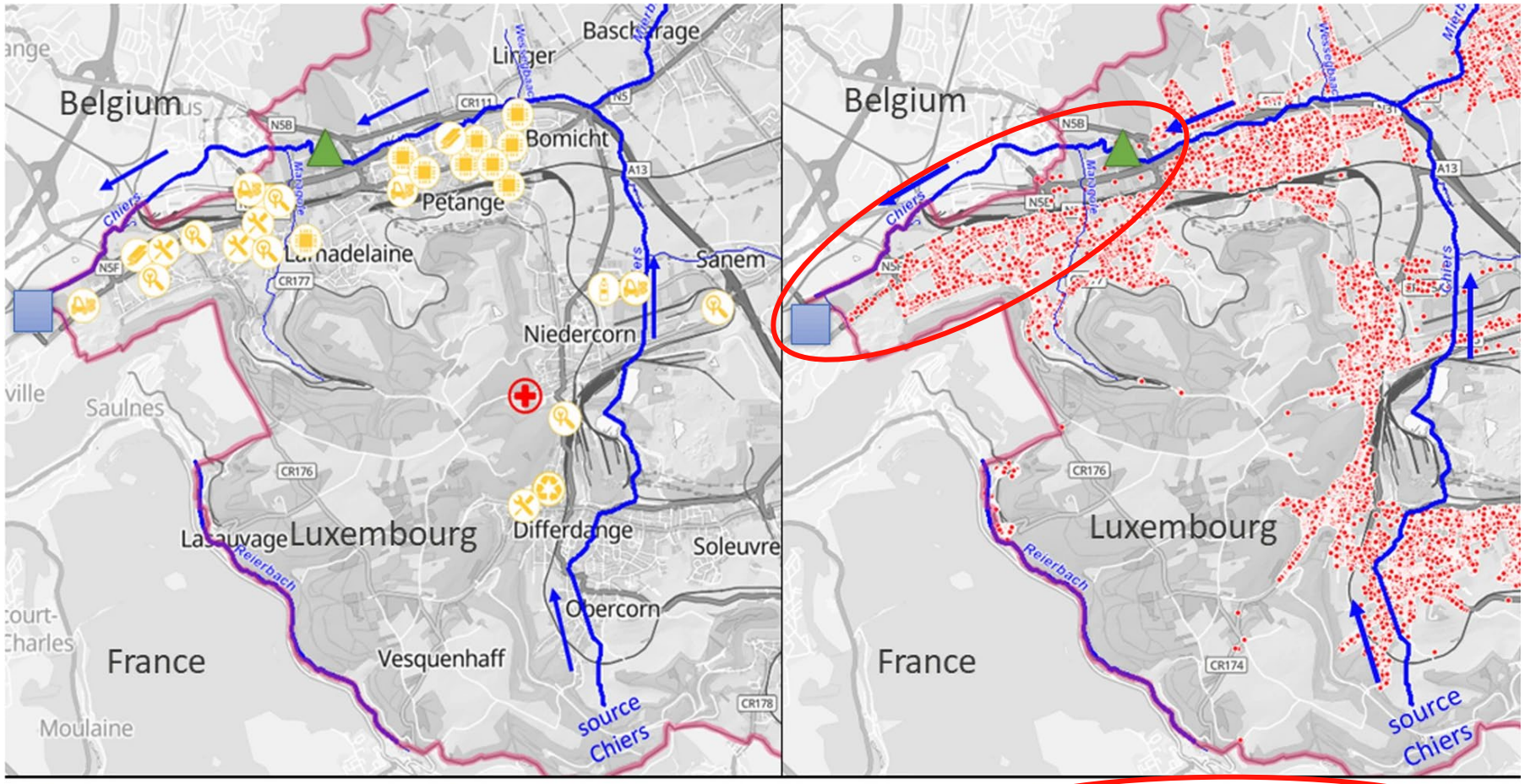


Temporal variations



Slide: D. Aurich, PhD Defense. Ref: Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

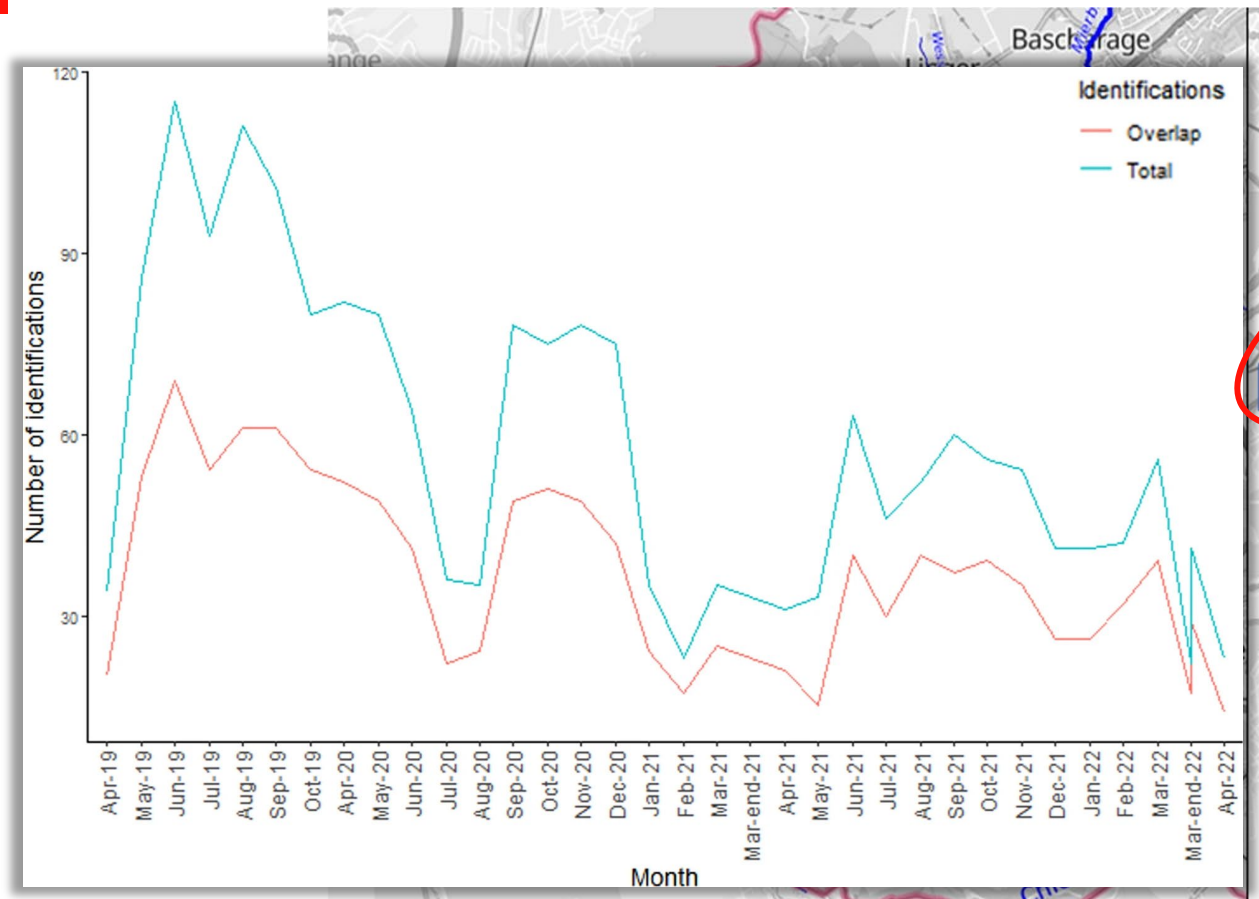
Spatial Assessment – Potential Chemical Sources at Chiers



WWTP Petange

Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

Spatial Assessment – Potential Chemical Sources at Chiers



WWTP Petange

Rubber goods & plastics
Logistics
Equipment & accessories
Recycling & maintenance
Metal: Equipment & accessories for metallurgy
Engineering
Electronics & electricity
Hospital
WWTP Petange (sampling point)
Sampling point AGE
Flow direction Chiers
Addresses: Populated Region

Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

Proposing New Candidates for Additional Monitoring

- 41 frequently-occurring chemicals not (yet) monitored

Pharmaceuticals

Industrial chemicals

Synonym	Use	Parent Name	PubChem CID	No of occurrences
Irbesartan	pharmaceutical	-	3749	32
Amisulpride	pharmaceutical	-	2159	31
Telmisartan	pharmaceutical	-	65999	26
Celiprolol	pharmaceutical	-	2663	25
Fluconazole	pharmaceutical	-	3365	25
Trimethoprim	pharmaceutical	-	5578	24
4-Acetamidoantipyrine	pharmaceutical	Metamizole	65743	23
4-NP	industrial	-	980	22
Desvenlafaxine	pharmaceutical	Venlafaxine	125017	22
TCEP	flame retardant	-	8295	19
TCPP	flame retardant	-	26176	19
Triethyl phosphate	industrial	-	6535	18
Adipic acid	industrial	-	196	18
4-Formylaminoantipyrine	pharmaceutical	Aminopyrine	72666	17
Carbamazepine-10,11-epoxide	pharmaceutical	Carbamazepine	2555	15
Dibutyl phthalate	industrial	-	3026	14
PFOA	industrial	-	9554	14
2-Hydroxycarbamazepine	pharmaceutical	Carbamazepine	129274	14
3-Hydroxypyridine	industrial	-	7971	13
Tributylamine	industrial	-	7622	13
D617	pharmaceutical	Verapamil	93168	13
Sulisobenzone	consumer products	-	19988	13
Ensulizole	consumer products	-	33919	13

Slide mod. from: D. Aurich, PhD Defense. Ref: Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

Proposing New Candidates for Additional Monitoring

- 41 frequently-occurring chemicals not (yet) monitored

Pharmaceuticals

Industrial chemicals

OP flame retardants

Synonym	Use	Parent Name	PubChem CID	No of occurrences
Irbesartan	pharmaceutical	-	3749	32
Amisulpride	pharmaceutical	-	2159	31
Telmisartan	pharmaceutical	-	65999	26
Celiprolol	pharmaceutical	-	2663	25
Fluconazole	pharmaceutical	-	3365	25
Trimethoprim	pharmaceutical	-	5578	24
4-Acetamidoantipyrine	pharmaceutical	Metamizole	65743	23
4-NP	industrial	-	980	22
Desvenlafaxine	pharmaceutical	Venlafaxine	125017	22
TCEP	flame retardant	-	8295	19
TCPP	flame retardant	-	26176	19
Triethyl phosphate	industrial	-	6535	18
Adipic acid	industrial	-	196	18
4-Formylaminoantipyrine	pharmaceutical	Aminopyrine	72666	17
Carbamazepine-10,11-epoxide	pharmaceutical	Carbamazepine	2555	15
Dibutyl phthalate	industrial	-	3026	14
PFOA	industrial	-	9554	14
2-Hydroxycarbamazepine	pharmaceutical	Carbamazepine	129274	14
3-Hydroxypyridine	industrial	-	7971	13
Tributylamine	industrial	-	7622	13
D617	pharmaceutical	Verapamil	93168	13
Sulisobenzone	consumer products	-	19988	13
Ensulizole	consumer products	-	33919	13

Slide mod. from: D. Aurich, PhD Defense. Ref: Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

Proposing New Candidates for Additional Monitoring

- 41 frequently-occurring chemicals not (yet) monitored

Pharmaceuticals

Industrial chemicals

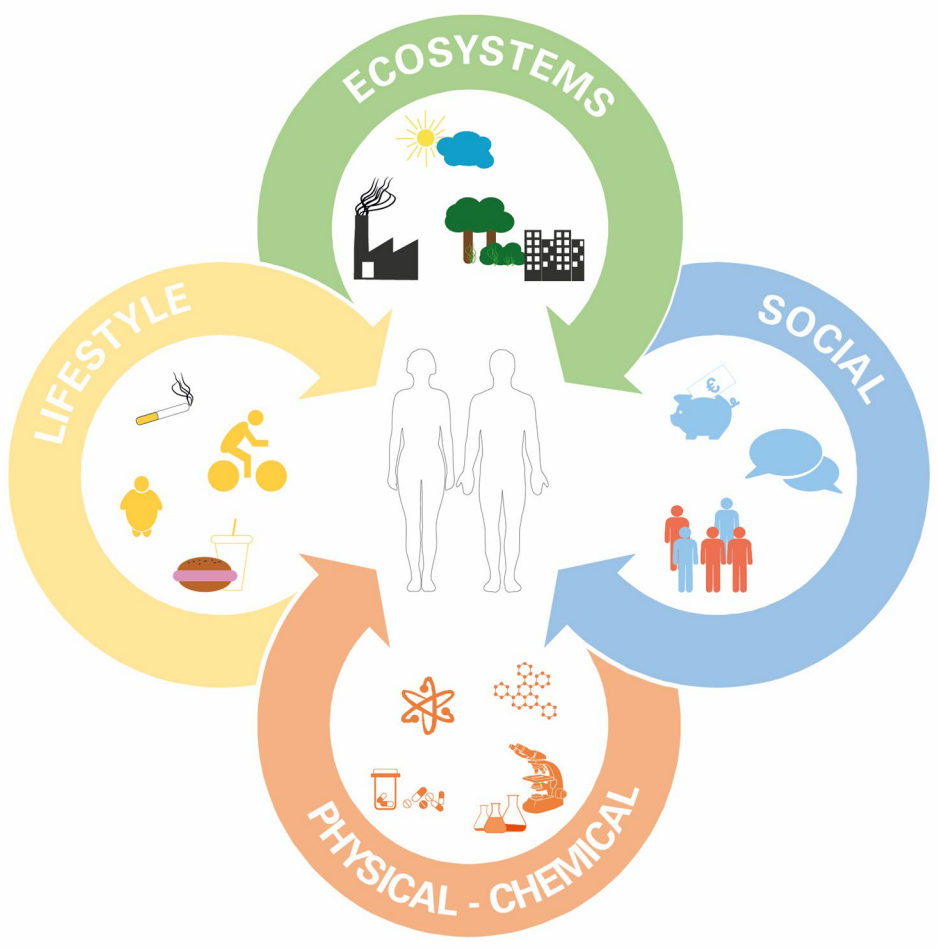
OP flame retardants

Transformation products

Synonym	Use	Parent Name	PubChem CID	No of occurrences
Irbesartan	pharmaceutical	-	3749	32
Amisulpride	pharmaceutical	-	2159	31
Telmisartan	pharmaceutical	-	65999	26
Celiprolol	pharmaceutical	-	2663	25
Fluconazole	pharmaceutical	-	3365	25
Trimethoprim	pharmaceutical	-	5578	24
4-Acetamidoantipyrine	pharmaceutical	Metamizole	65743	23
4-NP	industrial	-	980	22
Desvenlafaxine	pharmaceutical	Venlafaxine	125017	22
TCEP	flame retardant	-	8295	19
TCPP	flame retardant	-	26176	19
Triethyl phosphate	industrial	-	6535	18
Adipic acid	industrial	-	196	18
4-Formylaminoantipyrine	pharmaceutical	Aminopyrine	72666	17
Carbamazepine-10,11-epoxide	pharmaceutical	Carbamazepine	2555	15
Dibutyl phthalate	industrial	-	3026	14
PFOA	industrial	-	9554	14
2-Hydroxycarbamazepine	pharmaceutical	Carbamazepine	129274	14
3-Hydroxypyridine	industrial	-	7971	13
Tributylamine	industrial	-	7622	13
D617	pharmaceutical	Verapamil	93168	13
Sulisobenzone	consumer products	-	19988	13
Ensulizole	consumer products	-	33919	13

Slide mod. from: D. Aurich, PhD Defense. Ref: Aurich *et al* (2023) Env. Sci. Europe. DOI: 10.1186/s12302-023-00805-5

The Exposome meets the Metabolome

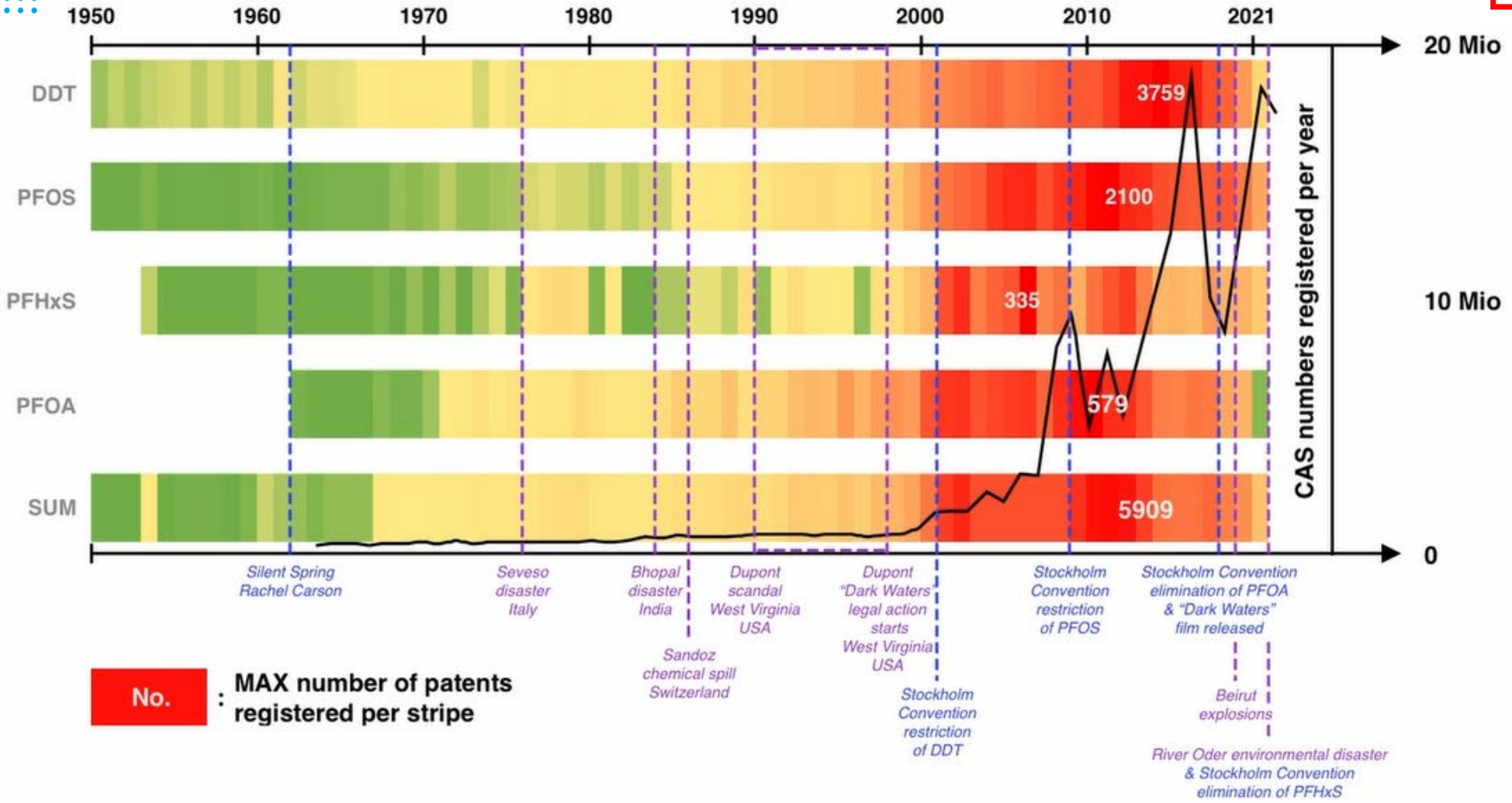


"The cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behaviour and endogenous processes"
(Miller & Jones, 2014)

Mod. from Vermeulen, Schymanski *et al.* (2020). Science. DOI: 10.1126/science.aay3164, Miller & Jones (2014) DOI: 10.1093/toxsci/kft251.

Our Chemical Past, Present, and Future

low / medium / high number of patents registered per year (WIPO)



Composition & Animation: Jamie Perera; Figure: Dagny Aurich; mod. from Arp et al. (2023) DOI: 10.1021/acs.est.3c01735. Credits: Jamie Perera, Dagny Aurich, Paul Thiessen, Emma Schymanski

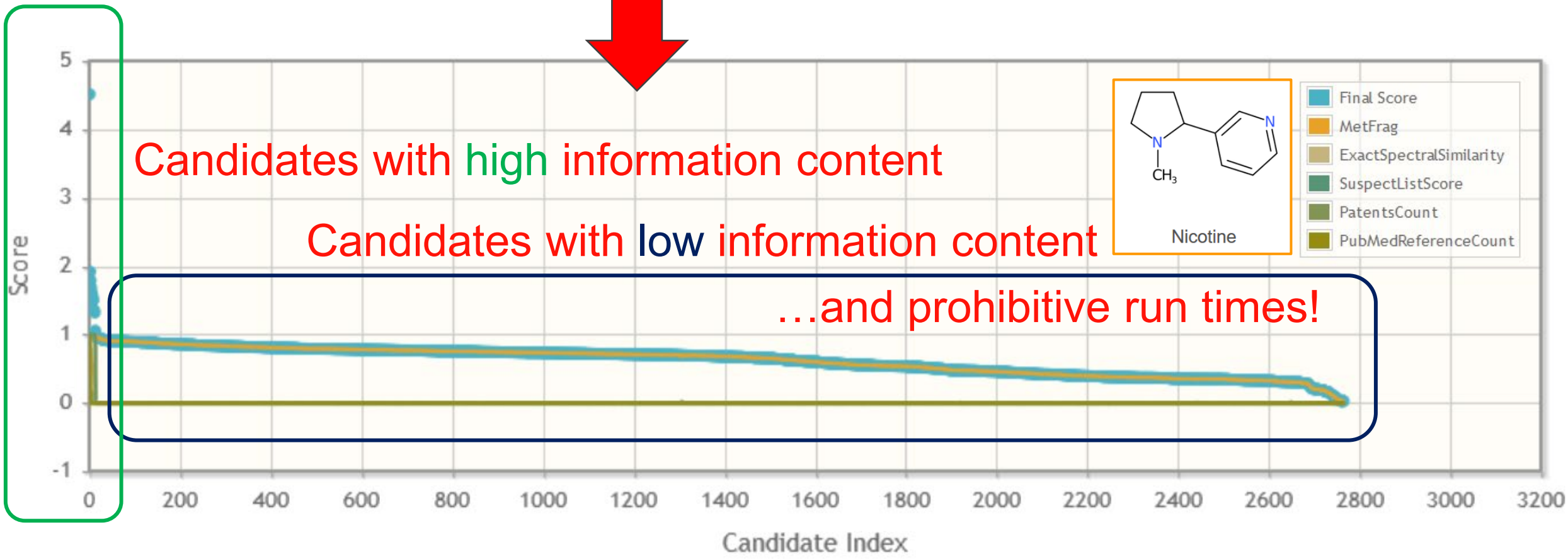
Grand Challenge: HR-MS “Chemical Space” is too Big!

CAS® A DIVISION OF THE AMERICAN CHEMICAL SOCIETY — 219 million

PubChem — 119 million

ChemSpider Search and share chemistry — 129 million

CompTox Chemicals Dashboard — 1.2 million



The Vast Chemical Space of Exposomics …

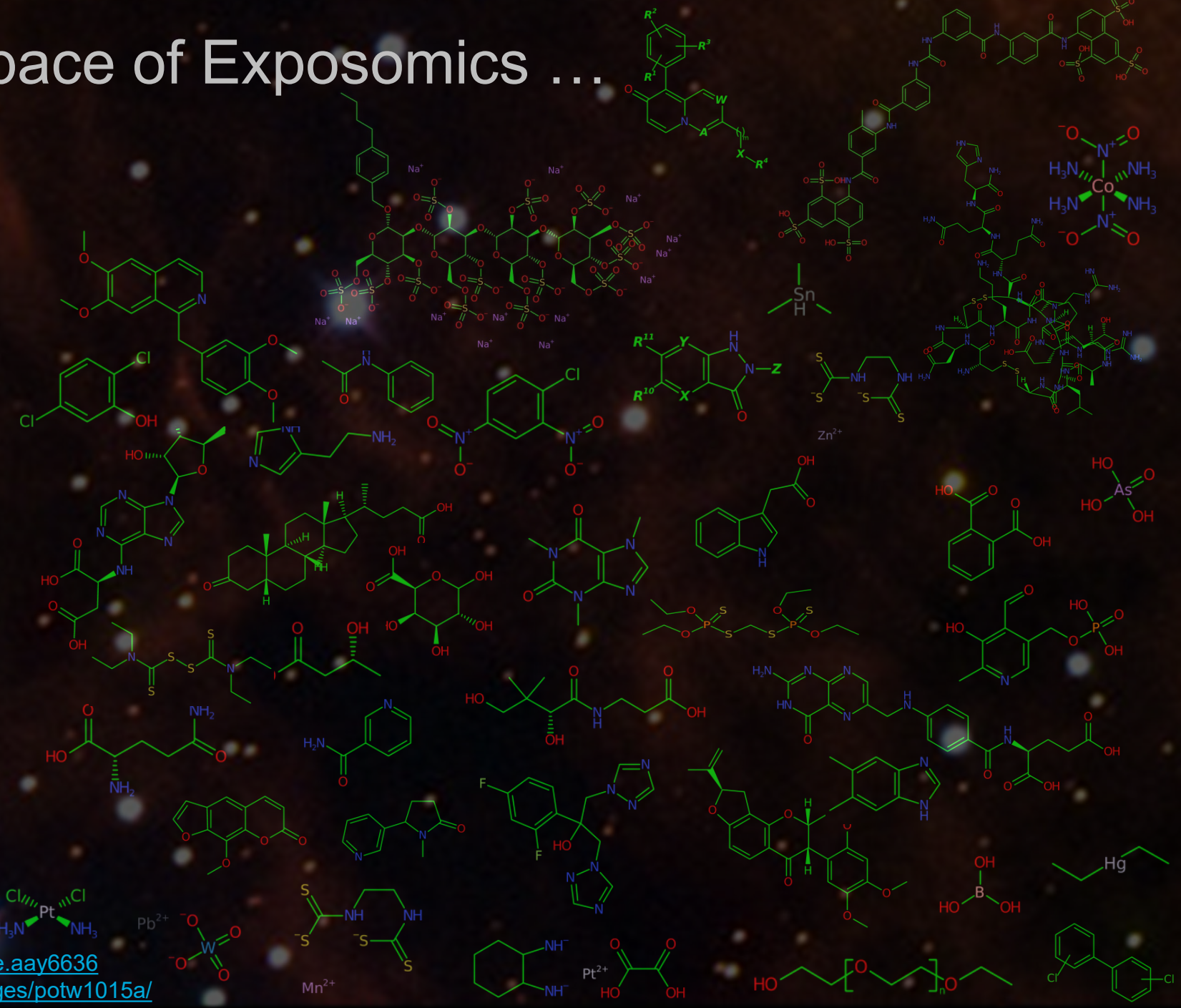
PubChem

119 million

MassBank

High Quality Mass Spectral Database

0.0078% of PubChem



Numbers updated 11 Sept. 2024.

Mod. from Escher *et al* (2020). Science. DOI: 10.1126/science.aay6636

Image: ESO/IDA/Danish 1.5 m http://www.eso.org/public/images/potw1015a/

The Vast Chemical Space of Exposomics …

PubChem

119 million

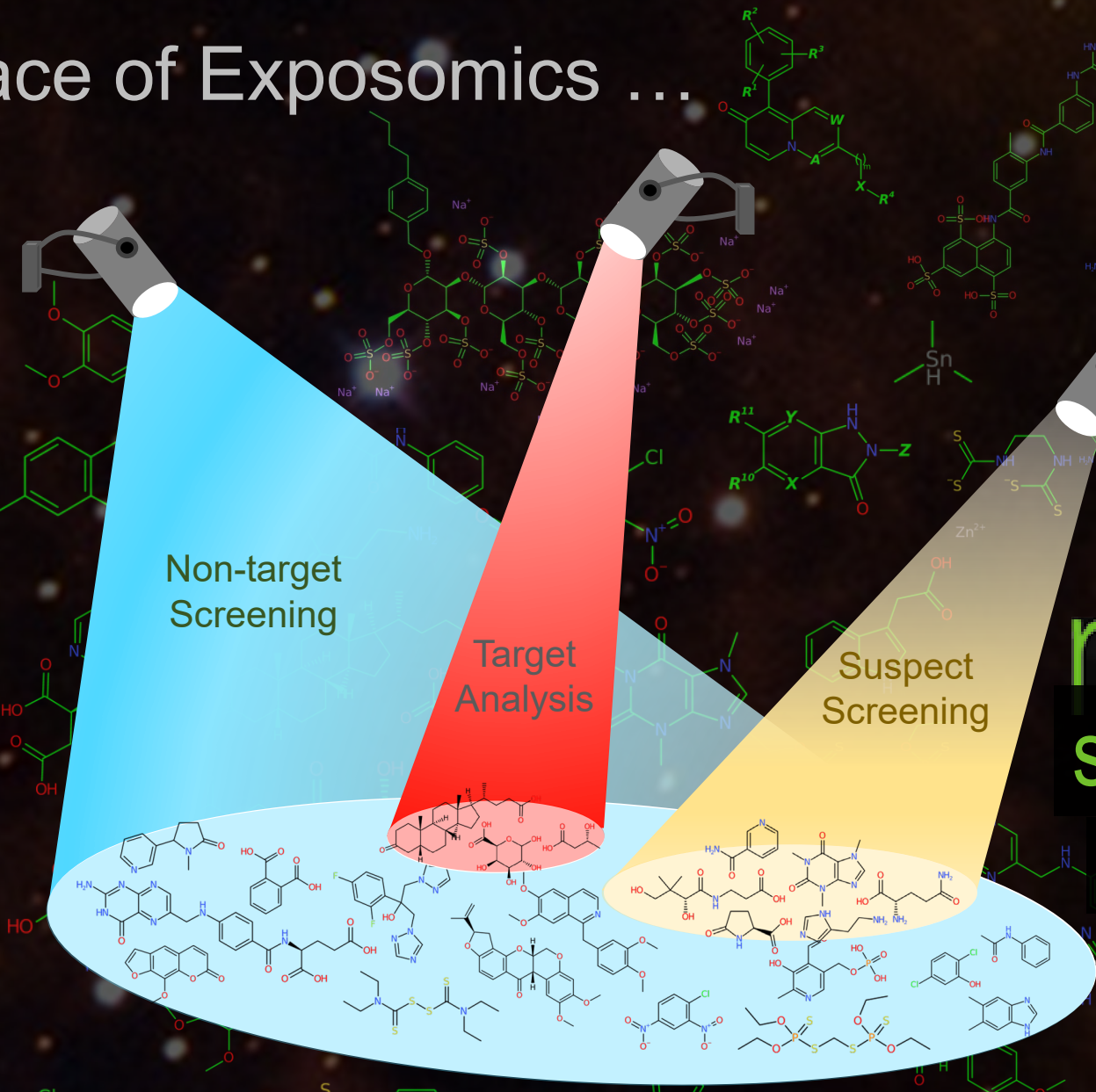
MassBank

High Quality Mass Spectral Database

0.0078% of PubChem

PubChemLite EXPOSOMICS

385,298



norman suspects

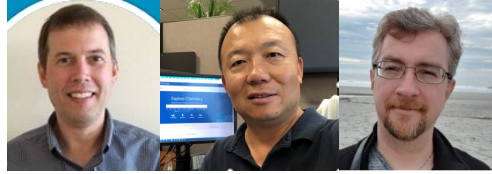
127,318

Numbers updated 11 Sept. 2024.

Mod. from Escher *et al* (2020). Science. DOI: 10.1126/science.aay6636

Image: ESO/IDA/Danish 1.5 m http://www.eso.org/public/images/potw1015a/

Subsetting PubChem



- PubChem Compound TOC 70,207,167
  - Agrochemical Information 3,145
  - Associated Disorders and Diseases 30,404
  - Biologic Description 2,521,405
  - Biological Test Results 4,687,906
  - Chemical and Physical Properties 323,548
  - Classification 23,893,099
  - Drug and Medication Information 20,284
  - Food Additives and Ingredients 62,272
  - Identification 4,616
  - Information Sources 49,281,547
  - Interactions and Pathways 214,389
  - Literature 2,326,701
  - Names and Identifiers 7,249,115
  - Patents 40,500,332
  - Pharmacology and Biochemistry 114,817
  - Related Records 13,741,310
  - Safety and Hazards 239,670
  - Spectral Information 1,625,326
  - Structures 15,020,428
  - Toxicity 117,290
  - Use and Manufacturing 67,453

- Agrochemical Information 3,145
  - Agrochemical Category 1,962
  - Agrochemical Transformations 1,488
  - EU Pesticides Data 1,243
  - USDA Pesticide Data Program 781

- Associated Disorders and Diseases 30,404
  - Disease and References 13,386

- Interactions and Pathways 214,389
  - Molecular Imaging Information 1,894
  - Protein Bound 3D Structures 76,529
  - Chemical-Target Interactions 85,888
  - Drug-Drug Interactions 5,214
  - Drug-Food Interactions 1,655
  - Pathways 62,568

PubChem Furathiocarb (Compound)

7.3 EU Pesticides Data

Active Substance	furathiocarb
Status	Not approved [Reg. (EC) No 1107/2009]
Legislation	2002/2076

PubChem Nicotine (Compound)

14 Associated Disorders and Diseases

Showing 11 to 15 of 252 items Download

Disease	Evidence Type	Evidence PMID
Airway Obstruction	marker/mechanism	33082140
Alzheimer Disease	therapeutic	16627626 17135361

PubChem Acetylcarnitine (Compound)

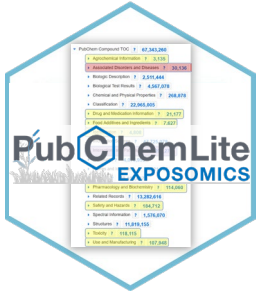
12.1 Pathways

4 items Download

Oxidation of Branched-Chain Fatty Acids
Source: PathBank External ID: SMP0000030
Taxonomy: Homo sapiens (human)

Schymanski *et al.* (2021) DOI: 10.1186/s13321-021-00489-0

PubChemLite for Exposomics



Home | Explore | About | Contact

PubChemLite EXPOSOMICS

Informative subset of PubChem relevant for various environmental, metabolomics, exposomics and mass spec. applications

Search

try C10H14N2 DUOANANYKYXIQY-UHFFFAOYSA-N atrazine

or Explore

What is PubChemLite?

PubChemLite is a subset of **384341** compounds from PubChem, with predicted collision cross section (CCS) values for **8** adducts.

PubChemLite exposomics is compiled from **10** categories:

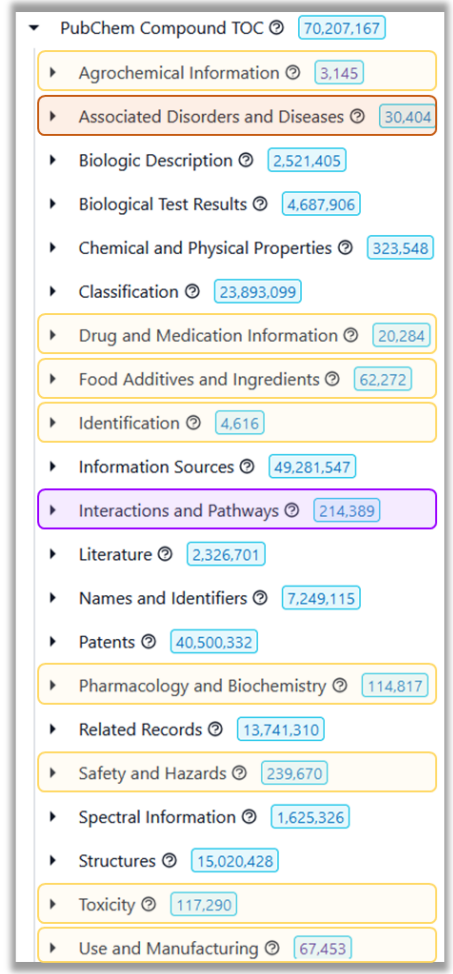
AgroChemInfo *BioPathway* *DrugMedicInfo*
FoodRelated *PharmacoInfo* *SafetyInfo*
ToxicityInfo *KnownUse* *DisorderDisease*
Identification

PubChemLite in a nutshell

PubChemLite is a subset of PubChem selected from major categories of the Table of Contents (TOC) page, including predicted collision cross section (CCS) values from CCSbase.

PubChemLite for Exposomics is currently compiled from 10 TOC categories: Agrochemical Information (AgroChemInfo), Associated Disorders and Diseases (DisorderDisease), Drug and Medication Information (DrugMedicInfo), Food Additives and Ingredients (FoodRelated), Identification (Identification), Interactions and Pathways – Pathways subset (BioPathway), Pharmacology and Biochemistry (PharmacoInfo), Safety and Hazards (SafetyInfo), Toxicity (ToxicityInfo), Use and Manufacturing (KnownUse).

CCS adducts provided from CCSbase are: [M+H]+, [M+K]+, [M+NH4]+, [M+Na-2H]-, [M+Na]+, [M-H]-, [M]+, [M]-.



http://pubchemlite.lcsb.uni.lu/

PubChemLite EXPOSOMICS

Search

try C10H14N2 DUOANANYKYXIQY-UHFFFAOYSA-N atrazine

or Explore

Informative subset of PubChem relevant for various environmental, metabolomics, exposomics and mass spec. applications

Structural Information

Molecular Formula

$C_8H_{14}ClN_5$

SMILES

CCNC1=NC(=NC(=N1)Cl)NC(C)C

InChI

InChI=1S/C8H14ClN5/c1-4-10-7-12-6(9)13-8(14-7)11-5(2)3/h5H,4H2,1-3H3,(H2,10,11,12,13,14)

InChIKey

MXWJVTOOROXGIU-UHFFFAOYSA-N

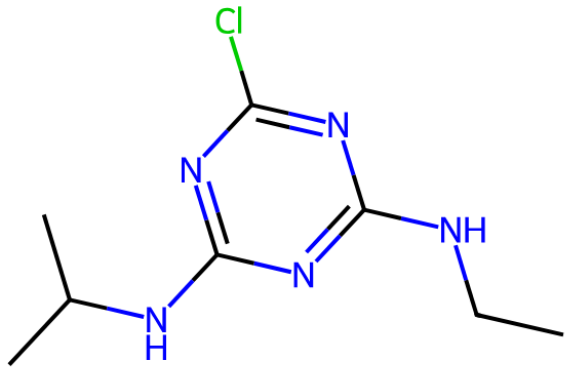
Compound name

6-chloro-4-N-ethyl-2-N-propan-2-yl-1,3,5-triazine-2,4-diamine

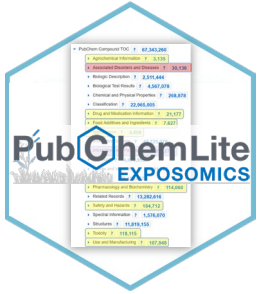
Related CIDs

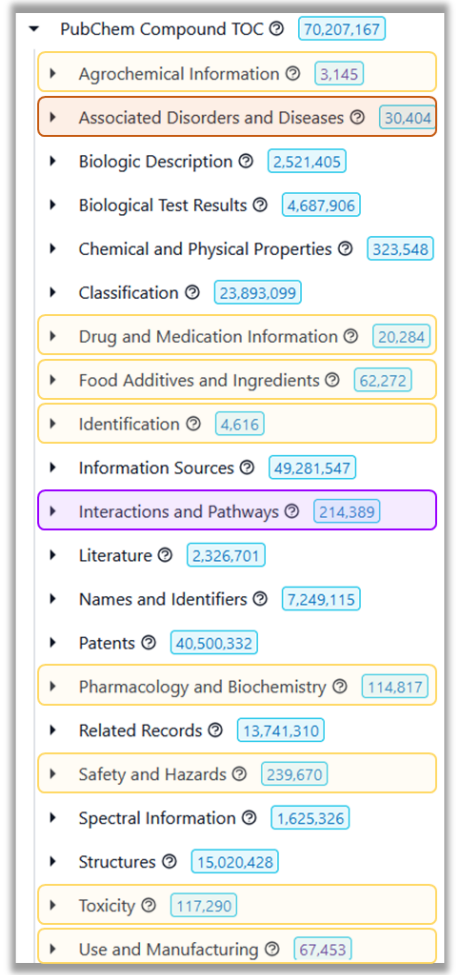
CID 2256 CID 12306645 CID 16213378

2D Structure



8	3616	51422	215.09378 Da	2.6
Annotation Hits	References	Patents	Monoisotopic Mass	XlogP (predicted)





http://pubchemlite.lcsb.uni.lu/e/compound/2256

PubChemLite EXPOSOMICS

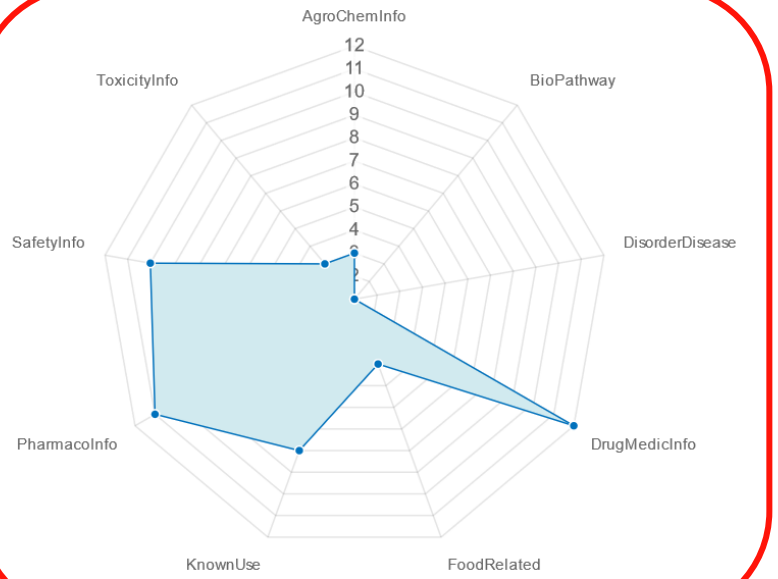
Search

19649

try C10H14N2 DUOANANYKYXIQY-UHFFFAOYSA-N atrazine

or Explore

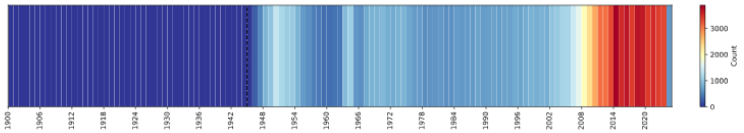
Informative subset of PubChem relevant for various environmental, metabolomics, exposomics and mass spec. applications



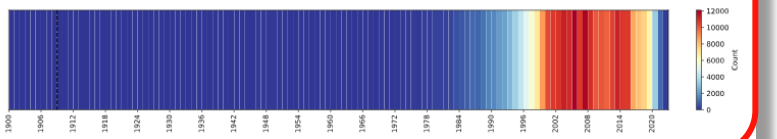
Adduct	m/z	Predicted CCS (Å²)
[M+H]+	582.27298	228.3
[M+Na]+	604.25492	229.2
[M+NH4]+	599.29952	230.1
[M+K]+	620.22886	228.0
[M-H]-	580.25842	222.5
[M+Na-2H]-	602.24037	243.9
[M]+	581.26515	228.2
[M]-	581.26625	228.2

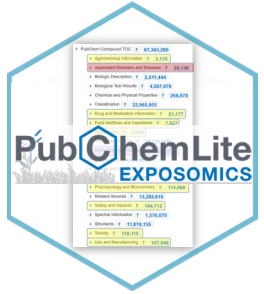
m/z: mass to charge ratio of the adduct.
Predicted Collision Cross Section (CCS) values (Å²) per adduct calculated using CCSbase.

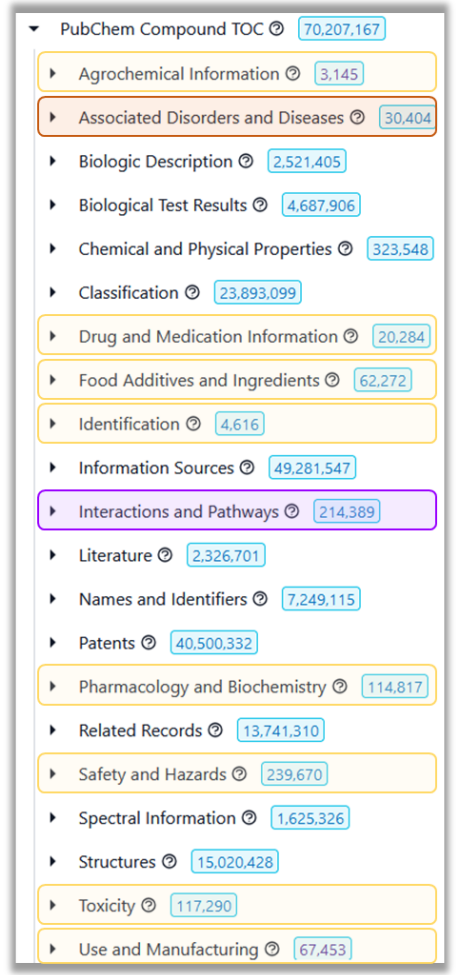
Literature stripe



Patent stripe



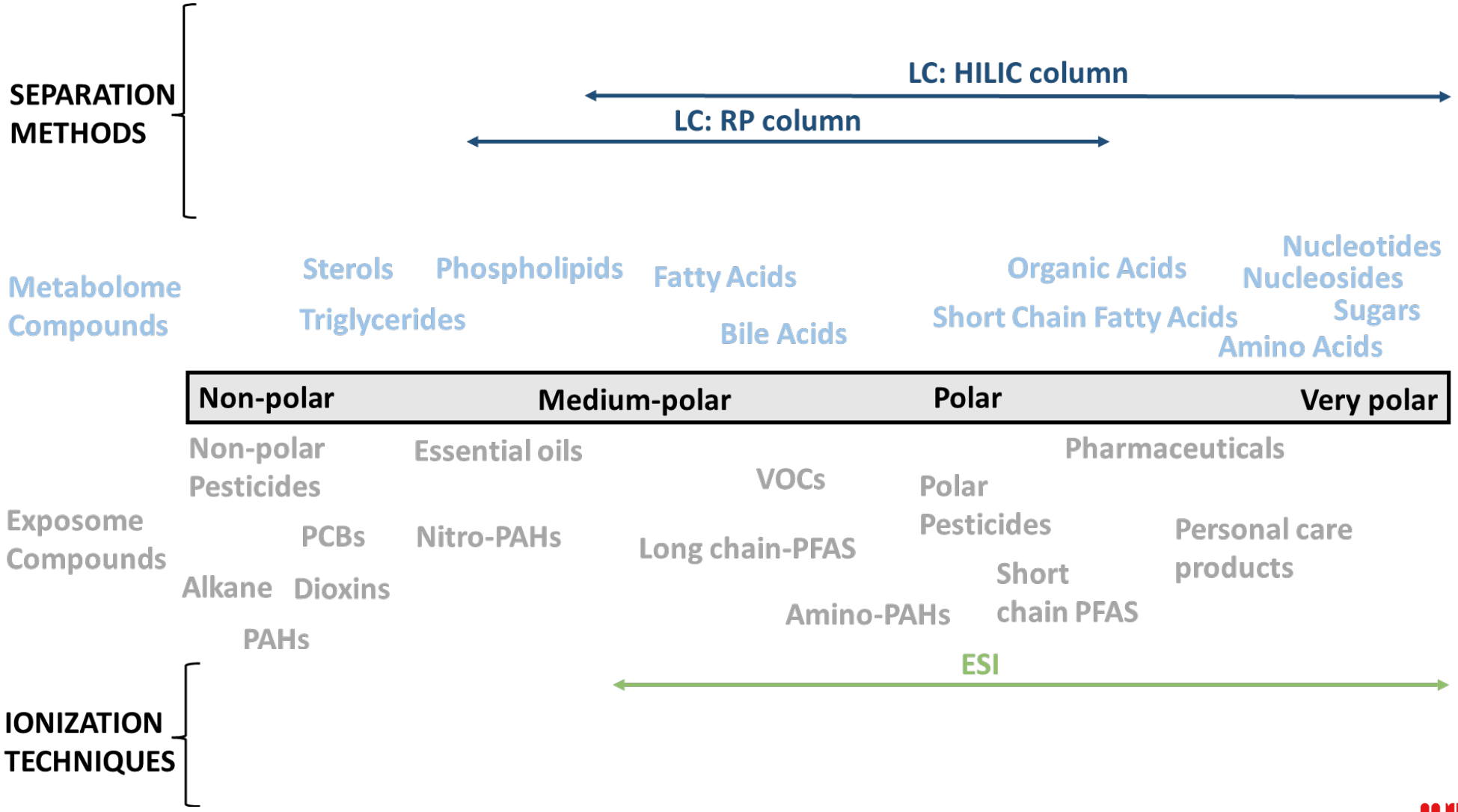




http://pubchemlite.lcsb.uni.lu/e/compound/19649

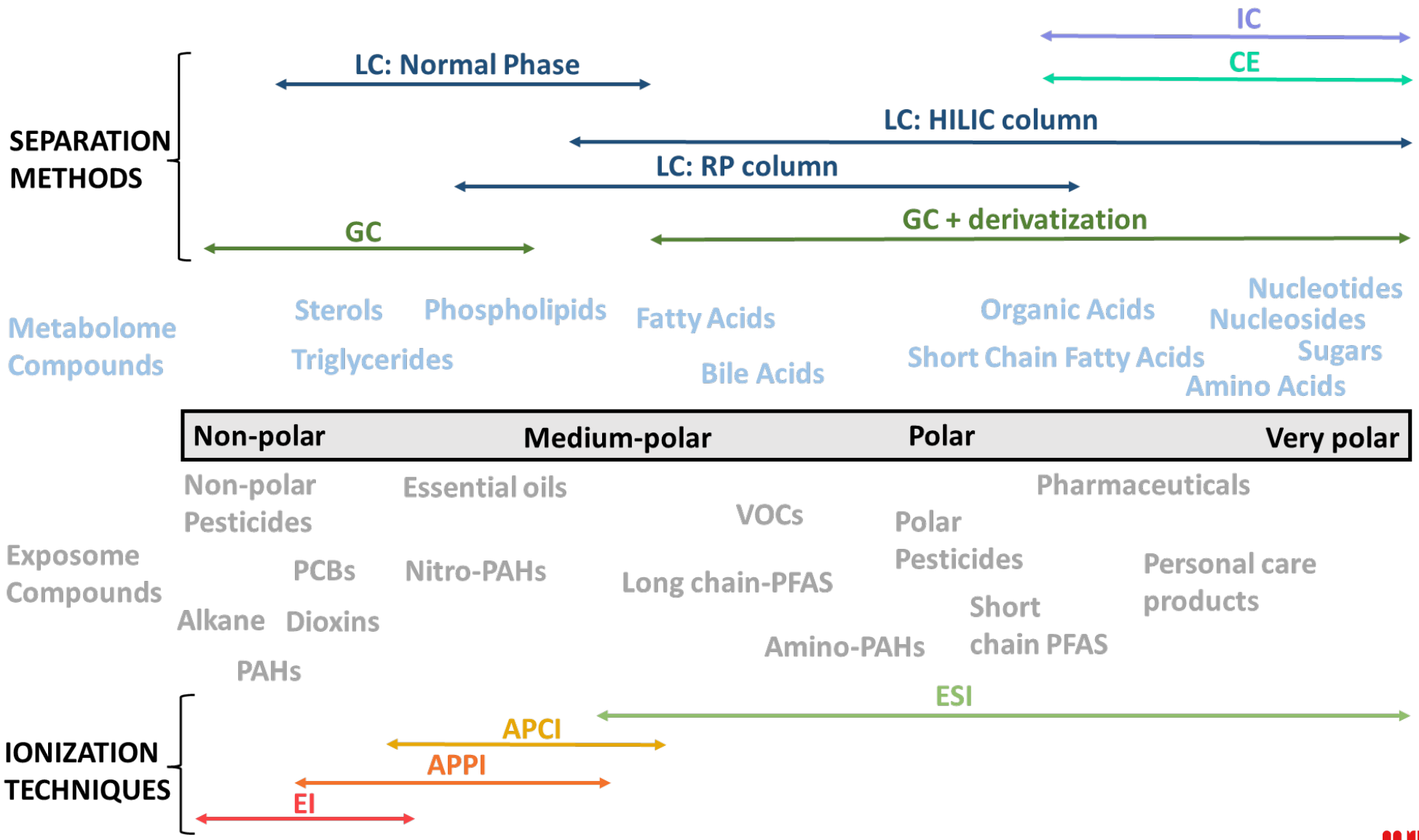
Analytical Challenge: Covering Multiple Chemical Classes…

MICROH



Source: Begoña Talavera Andújar, PhD Thesis.

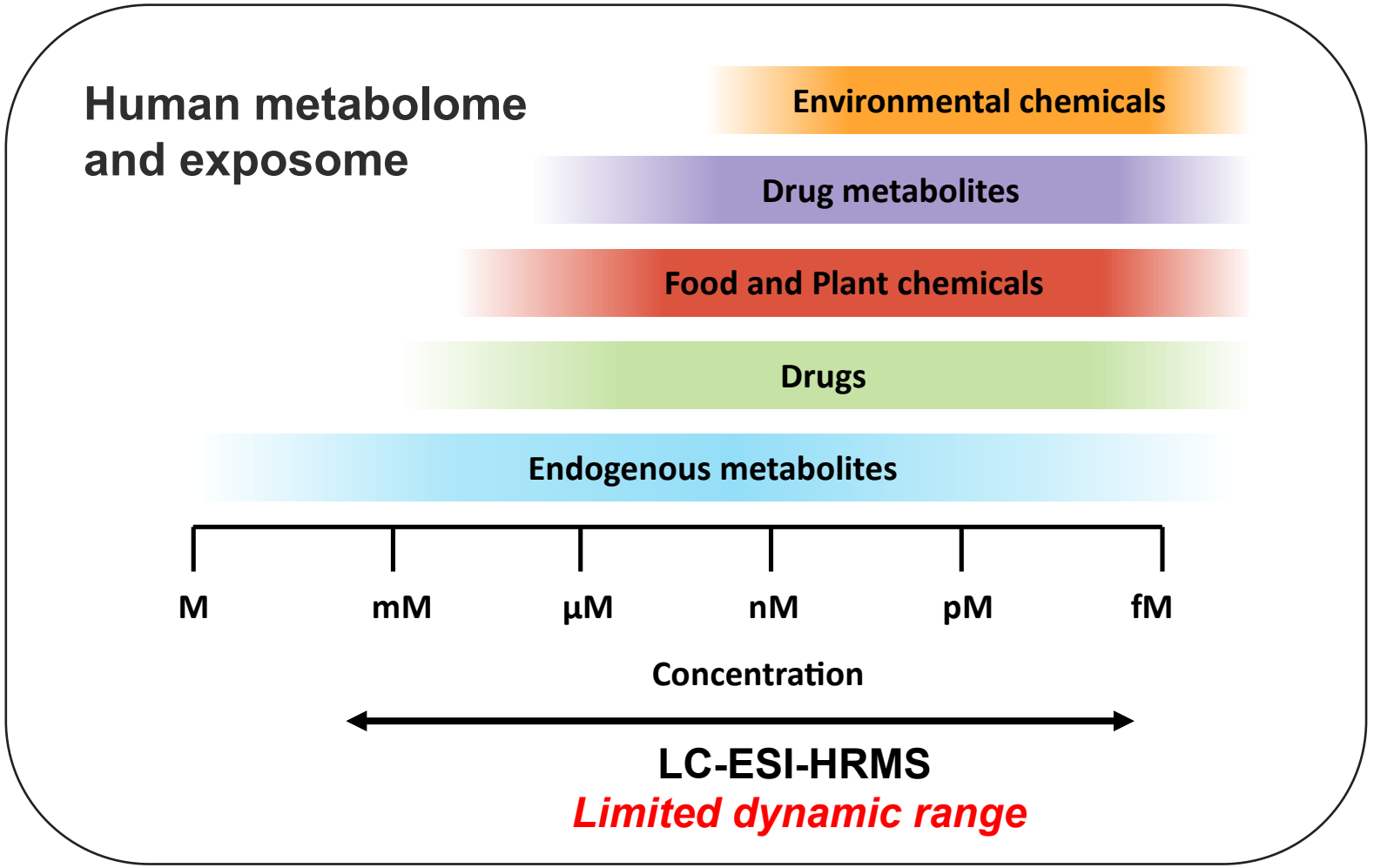
Analytical Challenge: Covering Multiple Chemical Classes…



Source: Begoña Talavera Andújar, PhD Thesis.

Analytical Challenge: Instrument Limitations …

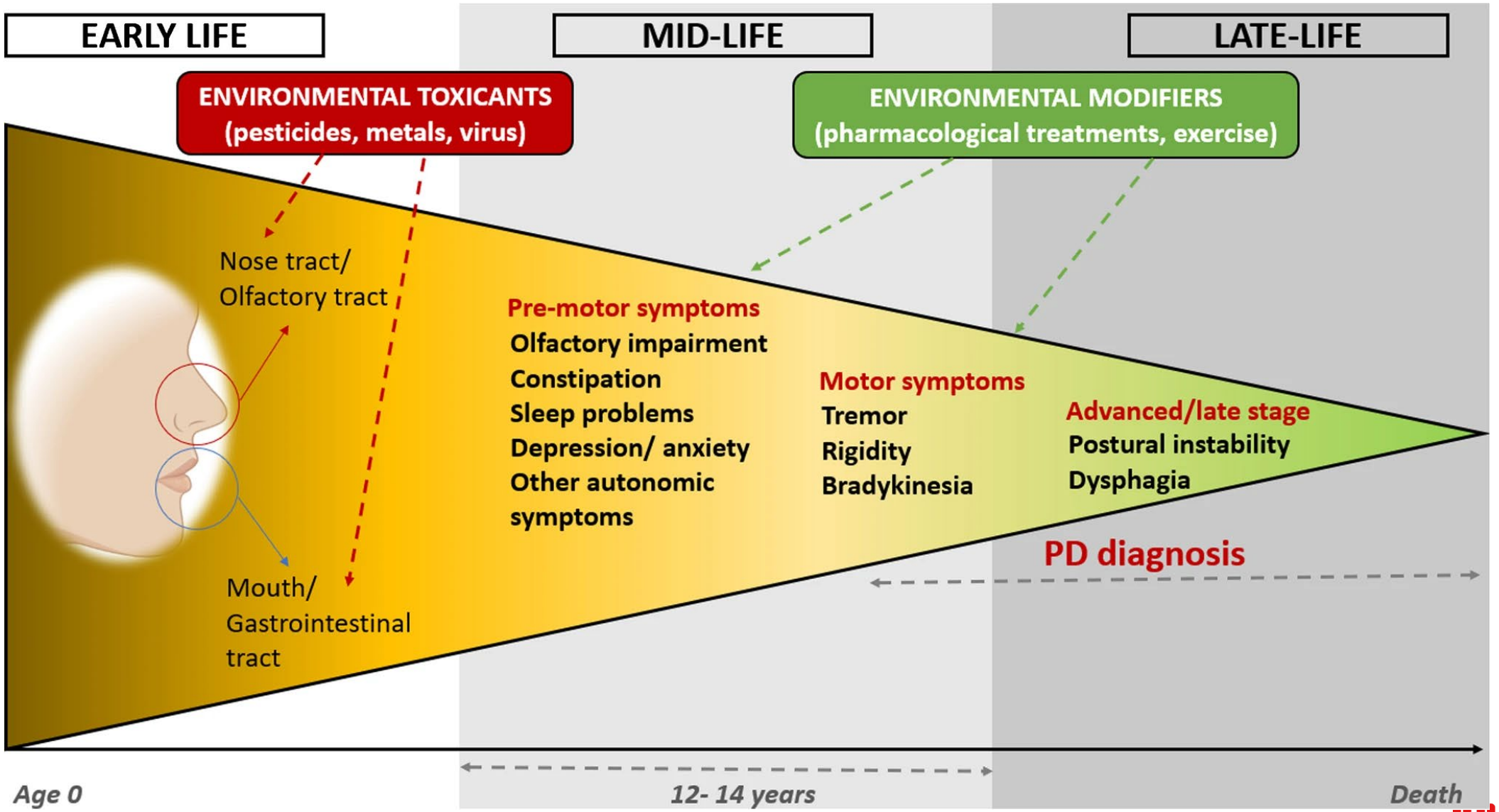
MICROH



Source: Begoña Talavera Andújar, PhD Thesis.

Timeline Challenge: Exposomics and Neurodegeneration

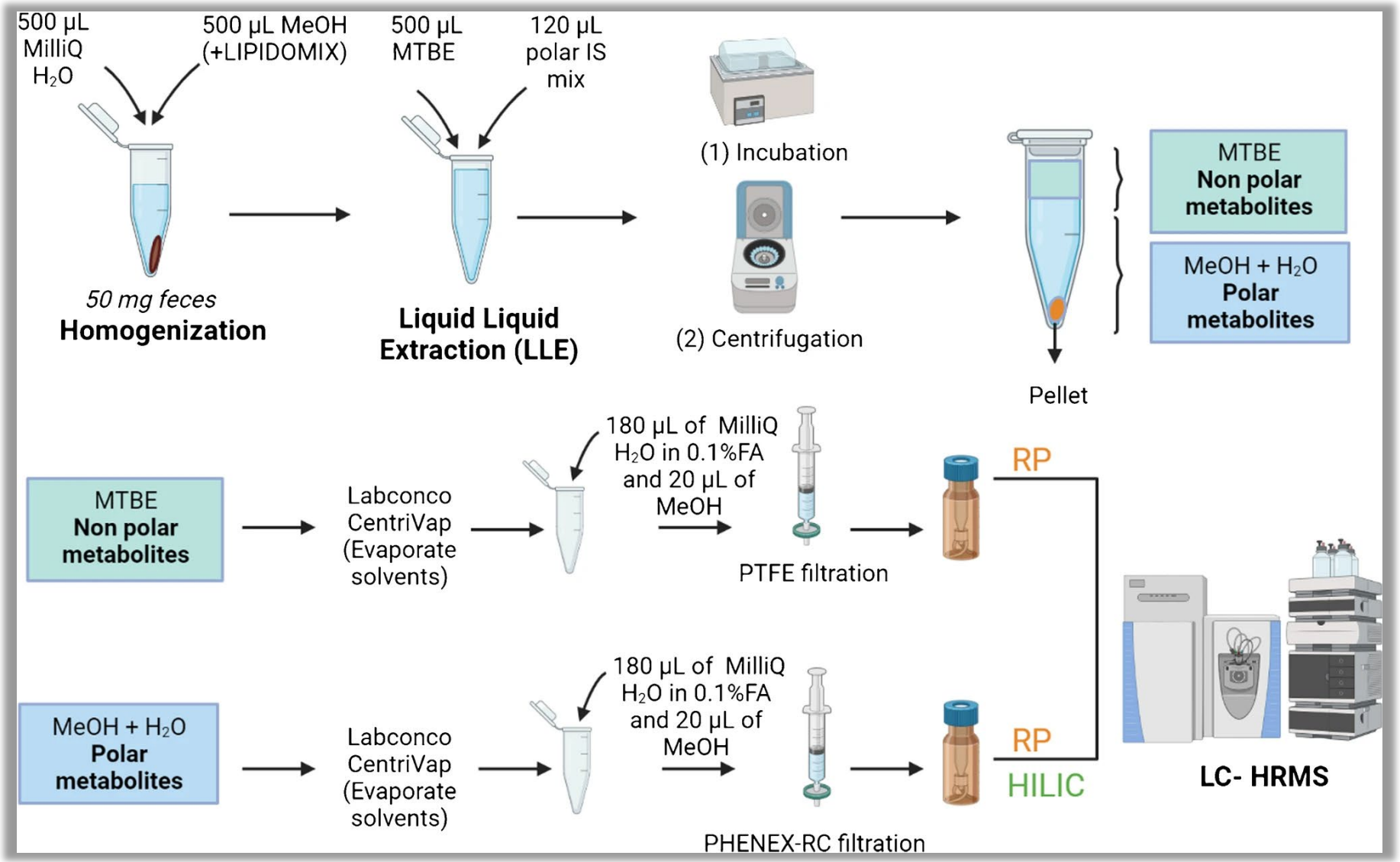
MICROH



Talavera Andújar B. et al. (2022). Anal Bioanal Chem. DOI: 10.1007/s00216-022-04207-z

Exposomics: Sample Prep. for Broad Range of Analysis

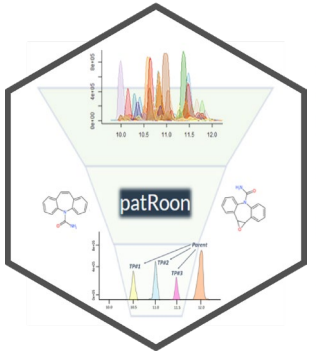
MICROH

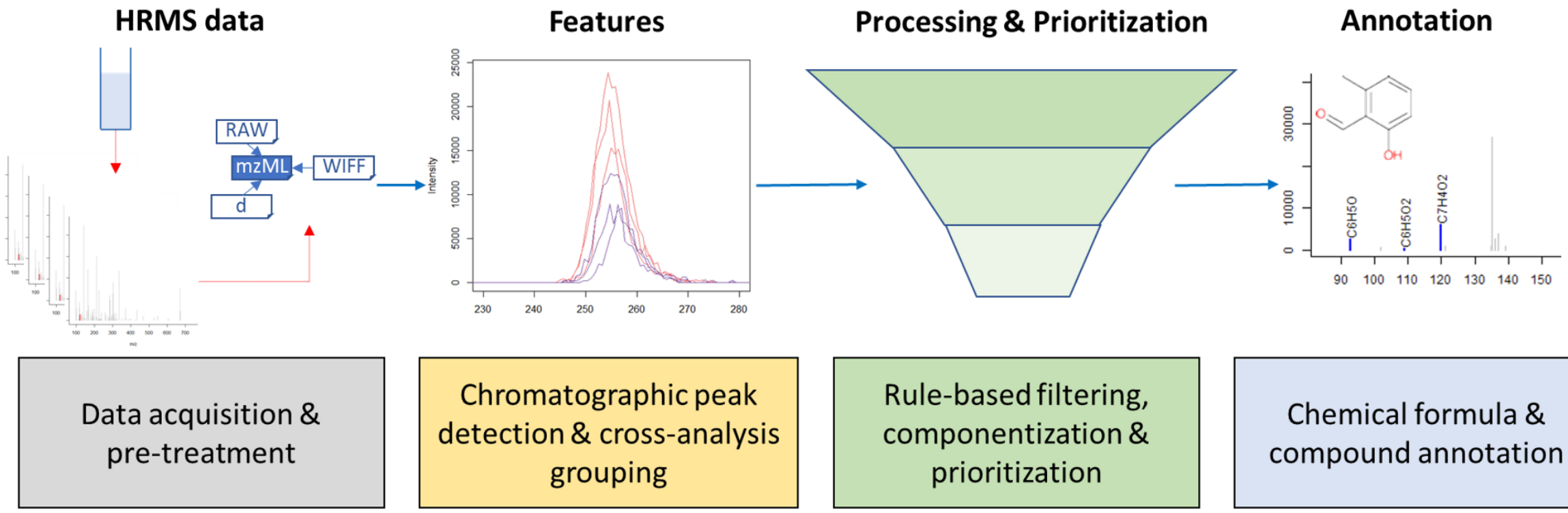


Talavera Andújar B. et al. (2022). Anal Bioanal Chem. DOI: 10.1007/s00216-022-04207-z

Open Source Workflows for NT-HRMS: patRoon

https://rickhelmus.github.io/patRoon/

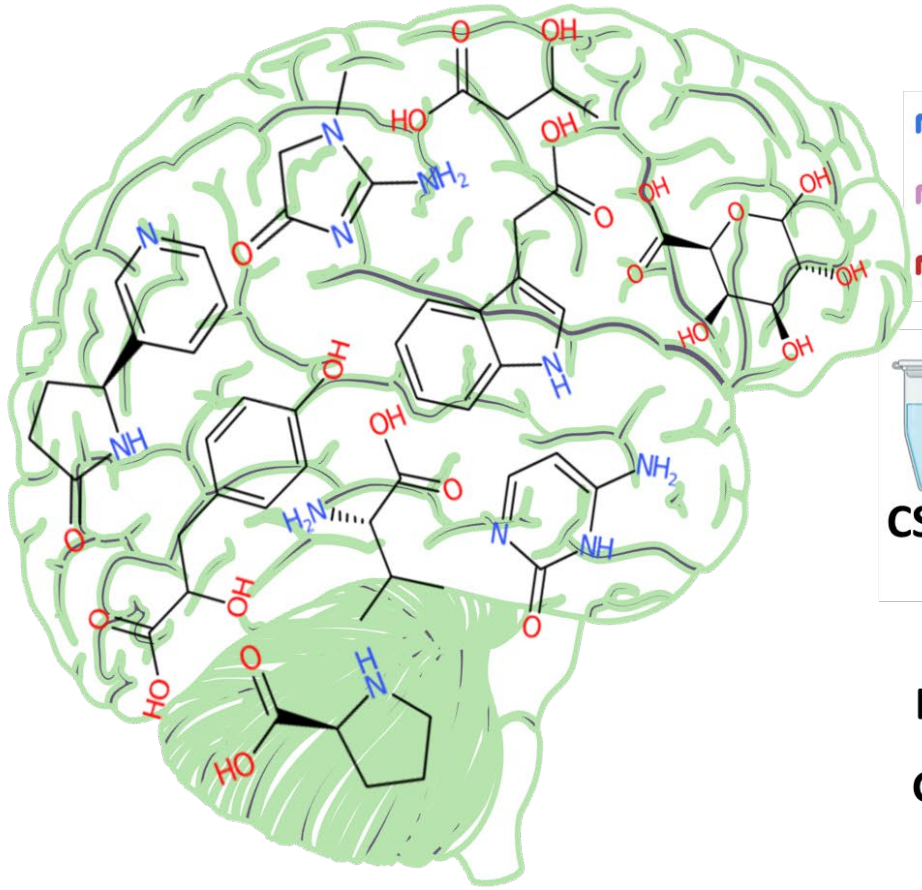


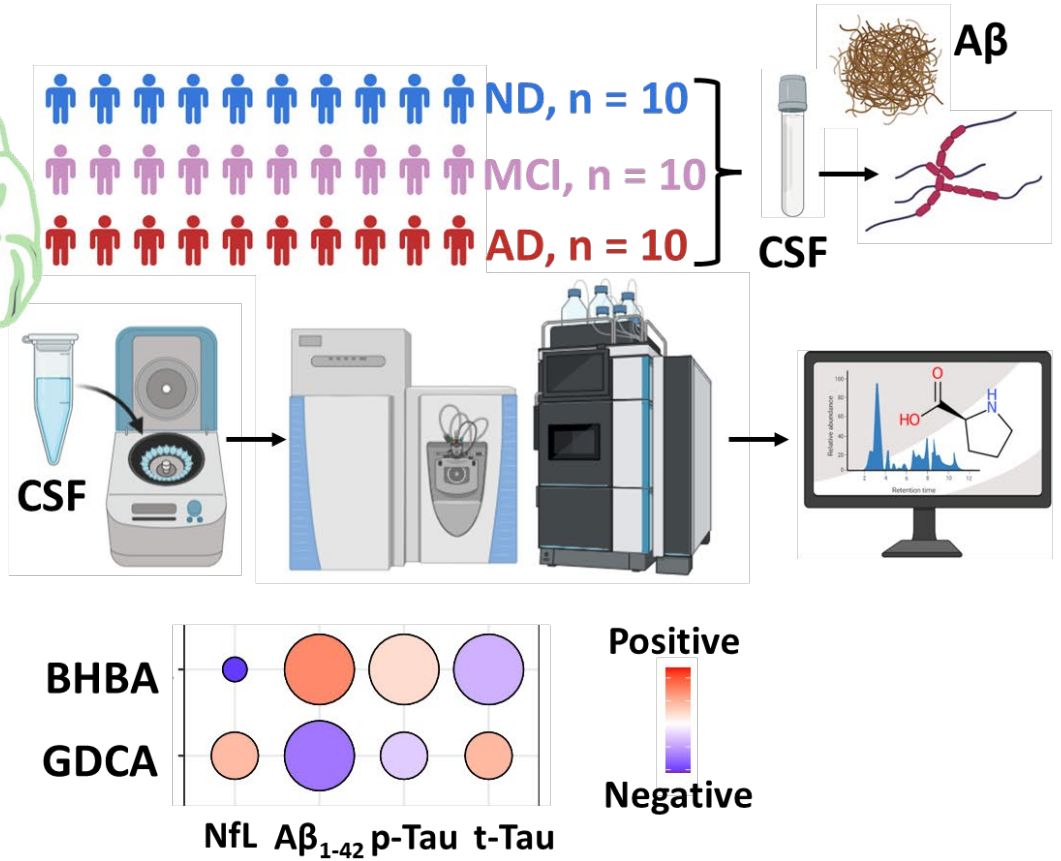


Helmus *et al.* (2022) patRoon 2.0 *J. Open Source Software,* DOI: 10.21105/joss.04029
Helmus *et al.* (2021) patRoon. *J. Cheminformatics*. DOI: 10.1186/s13321-020-00477-w

Exposomics Case Study (1): AD, MCI and CSF

MICROH

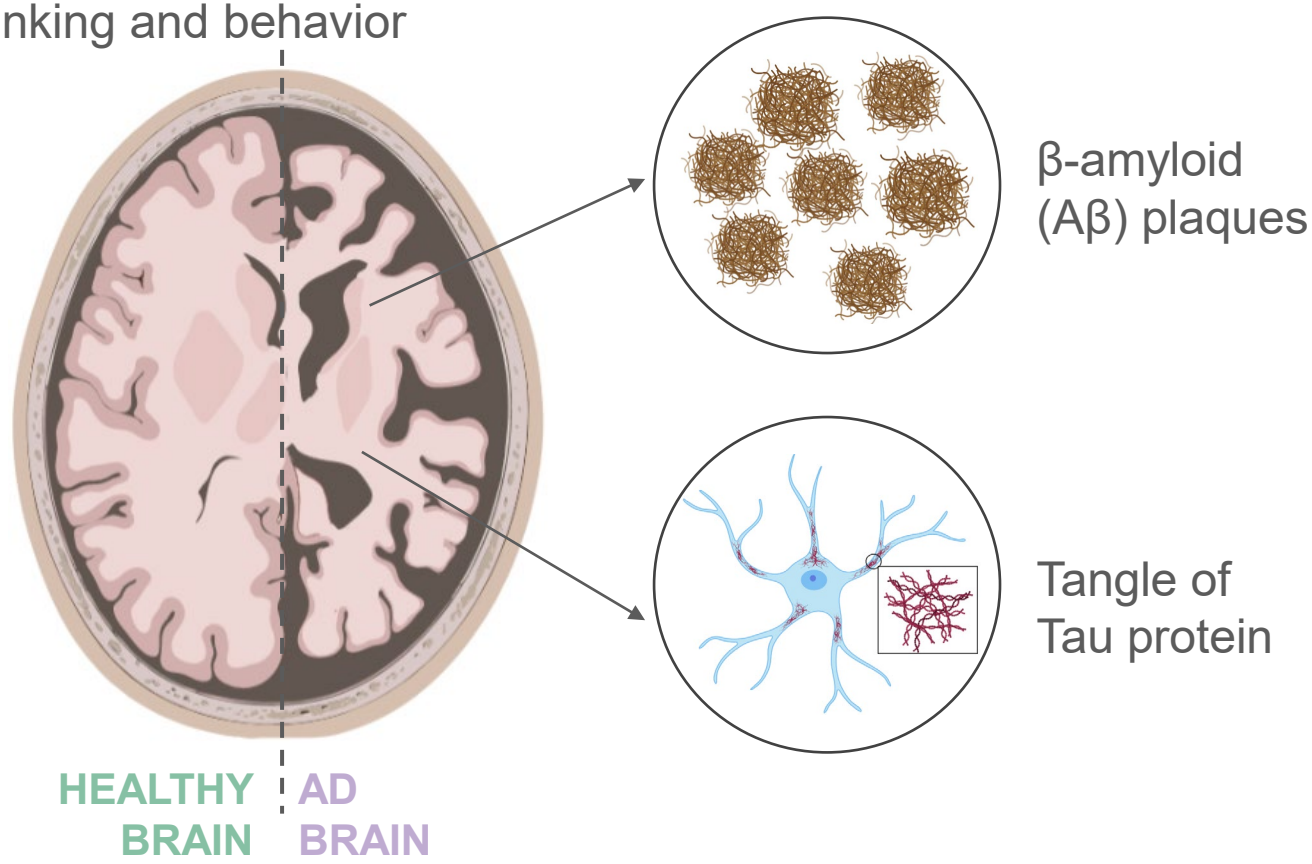




Talavera Andújar *et al*. (2024) ES&T, DOI: 10.1021/acs.est.3c10490

Alzheimer's Disease (AD)

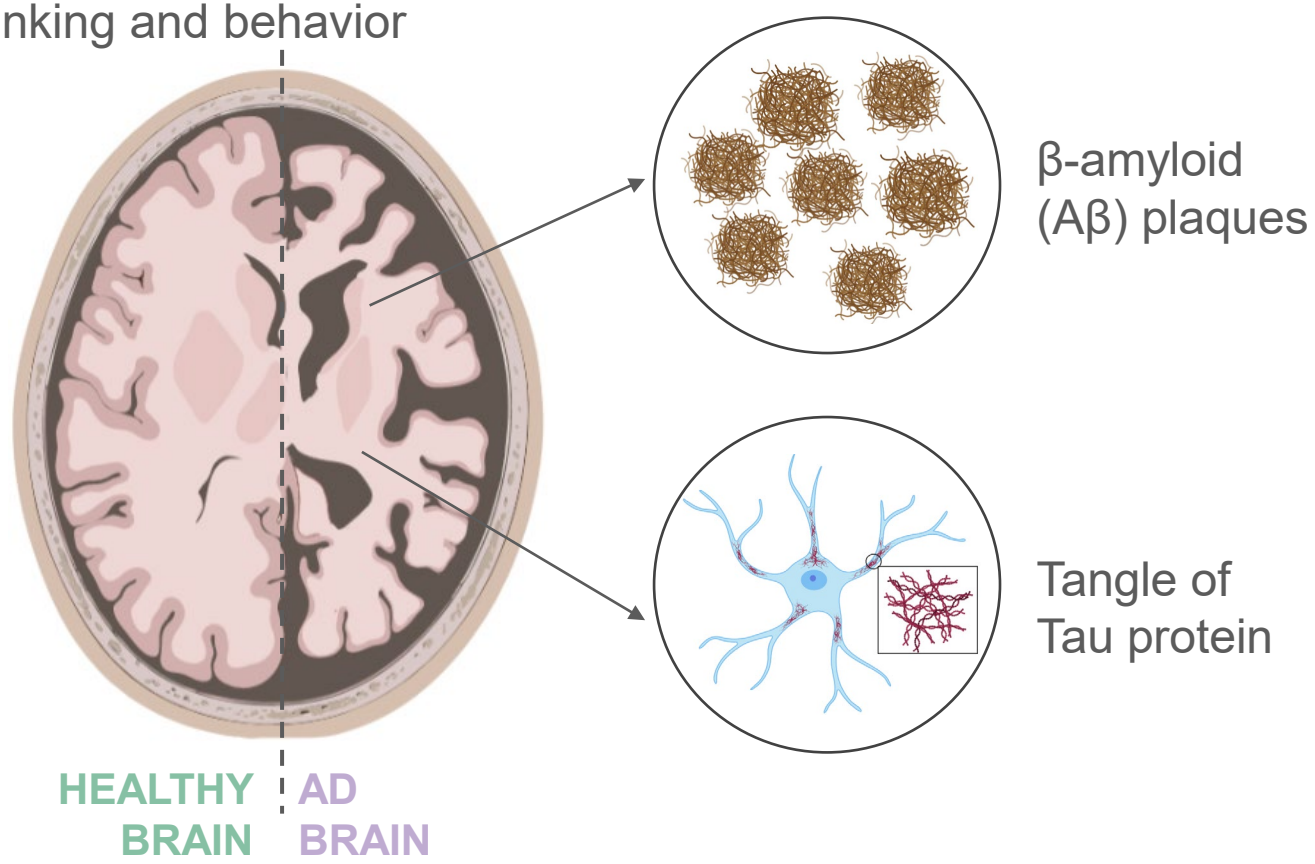
AD is a neurodegenerative disease that affects memory, thinking and behavior



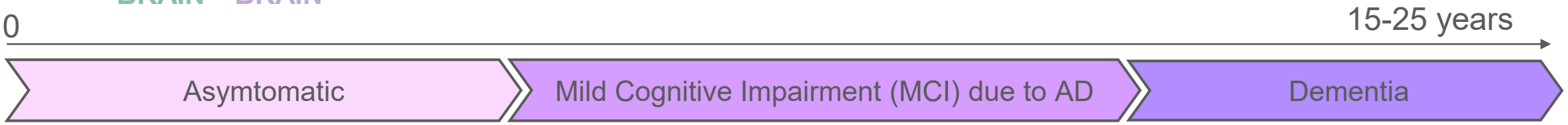
Source: Begoña Talavera Andújar. Refs: Hansson (2021). *Nat Medicine*. DOI: 10.1038/s41591-021-01382-x

Alzheimer's Disease (AD)

AD is a neurodegenerative disease that affects memory, thinking and behavior



AD is the most common cause of **dementia** accounting for **60-70 %** of dementia cases



Source: Begoña Talavera Andújar. Refs: Hansson (2021). *Nat Medicine*. DOI: 10.1038/s41591-021-01382-x

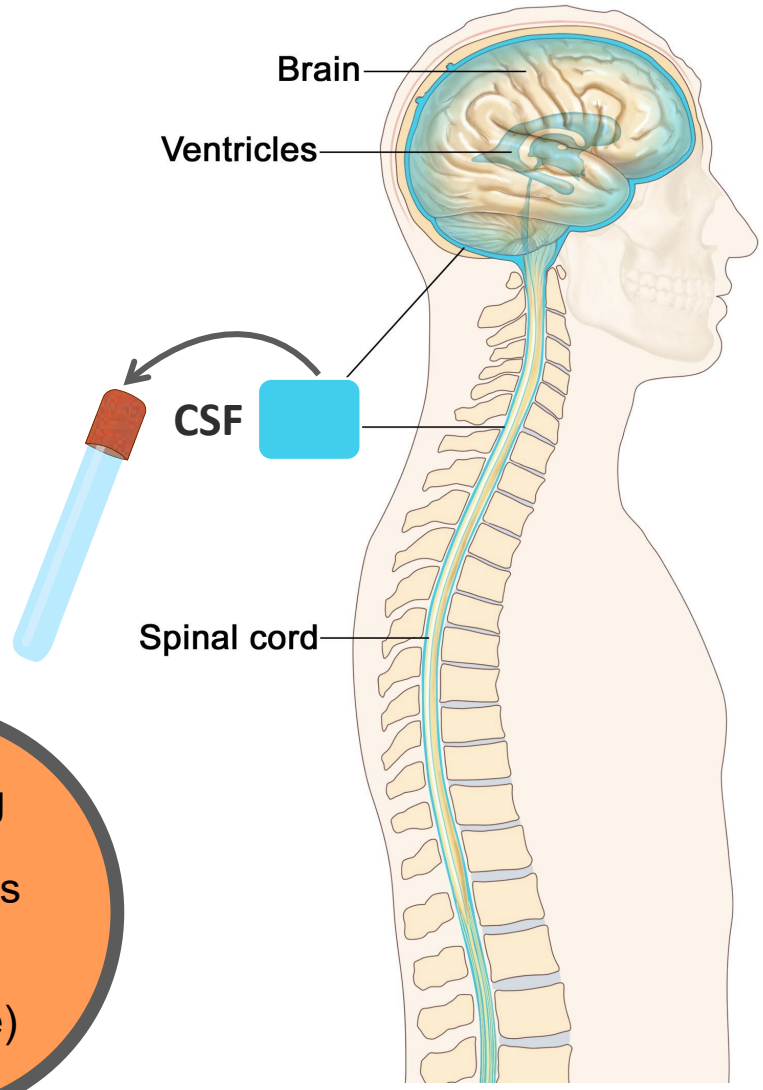
Cerebrospinal fluid (CSF) to study AD and MCI

AD

MCI

ND

CSF samples from three groups of subjects were collected and analyzed



Lack of CSF standard **reference material**

Limited exposomics/ metabolomics **studies** in CSF

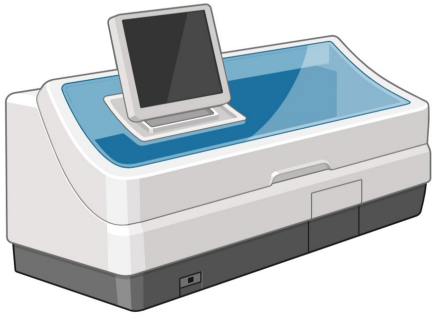
Most **exogenous** chemicals at **trace levels** compared to endogenous

Obtaining **control samples** is **difficult** (invasive procedure)

Source: Begoña Talavera Andújar & https://www.cancer.gov/publications/dictionaries/cancer-terms/def/cerebrospinal-fluid

Analytical Methods

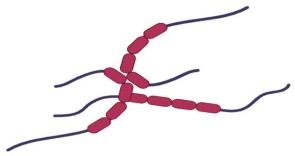
AD clinical biomarkers



Lumipulse G600II analyzer

Collaboration with
Neuroinflammation group (LCSB)

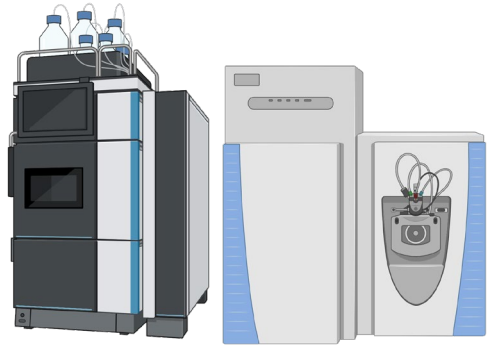
$A\beta_{40}$
$A\beta_{42}$



t-Tau
p-Tau

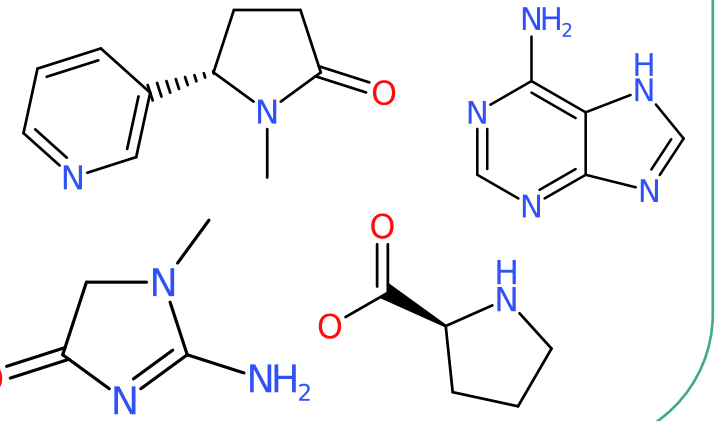
NfL

Non-target LC-HRMS

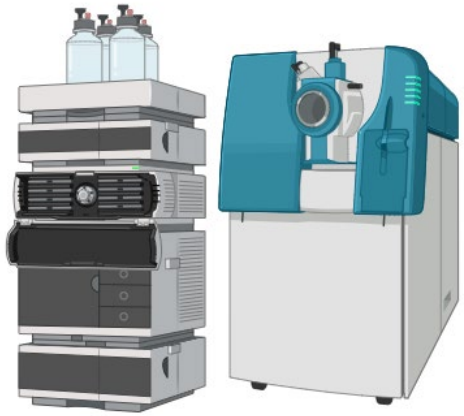


RP
HILIC

HRMS
Q Exactive™ HF



Target LC-MS (Bile Acids)



RP

Sciex 7500 QQQ

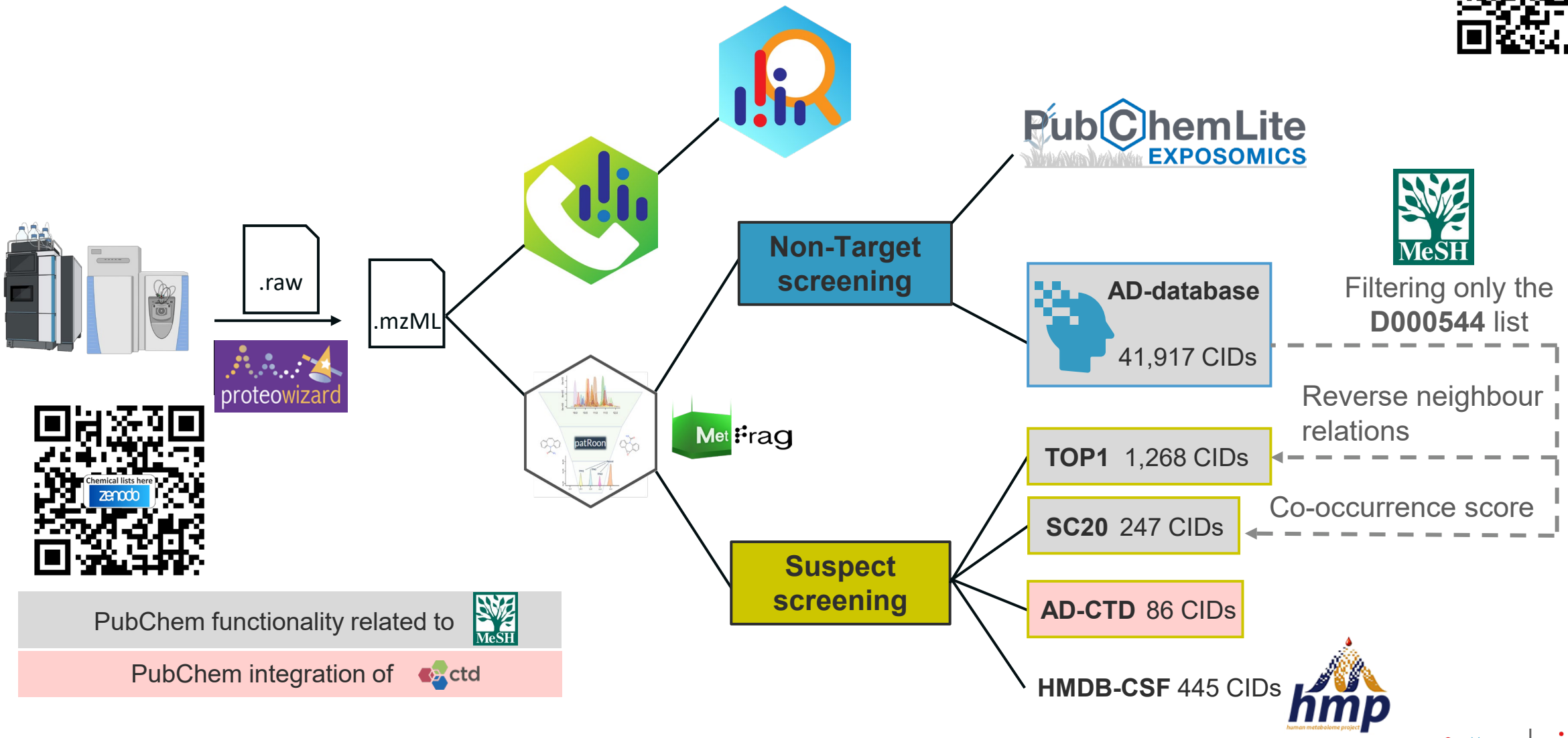
Analysis performed in Canada by:

UVic Genome BC PROTEOMICS CENTRE

Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

NfL: Neurofilament Light

Data Processing and Annotation



Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

Disease-specific Information in PubChem

PubChem Rotenone (Compound)

14 Associated Disorders and Diseases

83 items Download

Search

Disease	Evidence Type	Evidence PMID
Motor Disorders	marker/mechanism	26770656 27016191 29649621

PubChem Rotenone (Compound)

15.10 Chemical-Disease Co-Occurrences in Literature

Showing 3 of 25 View More Co-Occurrence and Evidence Data Download

Disease	Selected evidence
Parkinson Disease	1,590 articles View All Rotenone-Induced Model of Parkinson's Disease: Beyond Mitochondrial Complex I Inhibition PMID 36593435; DOI 10.1007/s12035-022-03193-8; Molecular neurobiology 2023 Jan; ?(?): (Review Article) Name matches: parkinson's disease *rotenone* HR LC-MS/MS metabolomic profiling of Yucca aloifolia fruit and the potential neuroprotective effect on rotenone-induced Parkinson's disease in rats PMID 36854038; DOI 10.1371/journal.pone.0282246; PloS one 2023; 18(2):e0282246 Name matches: parkinson's disease *rotenone*

Database

PubChem Protein, Gene, Pathway, and Taxonomy Data Collections: Bridging Biology and Chemistry through Target-Centric Views of PubChem Data

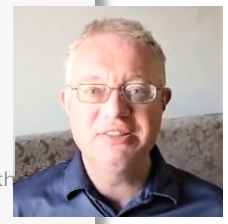
Sunghwan Kim †, Tiejun Cheng †, Siqian He, Paul A. Thiessen, Qingliang Li, Asta Gindulyte, Evan E. Bolton

MICROH

Discovering and Summarizing Relationships Between Chemicals, Genes, Proteins, and Diseases in PubChem

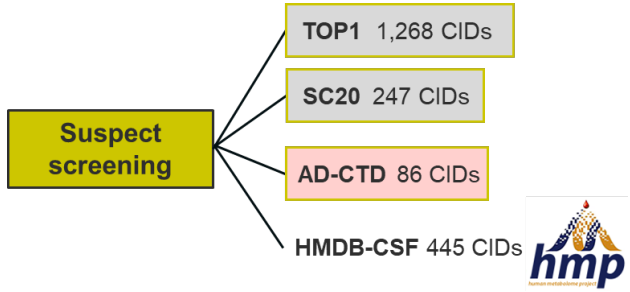
Leonid Zaslavsky*†, Tiejun Cheng†, Asta Gindulyte†, Siqian He†, Sunghwan Kim†, Qingliang Li†, Paul Thiessen†, Bo Yu† and Evan E. Bolton†

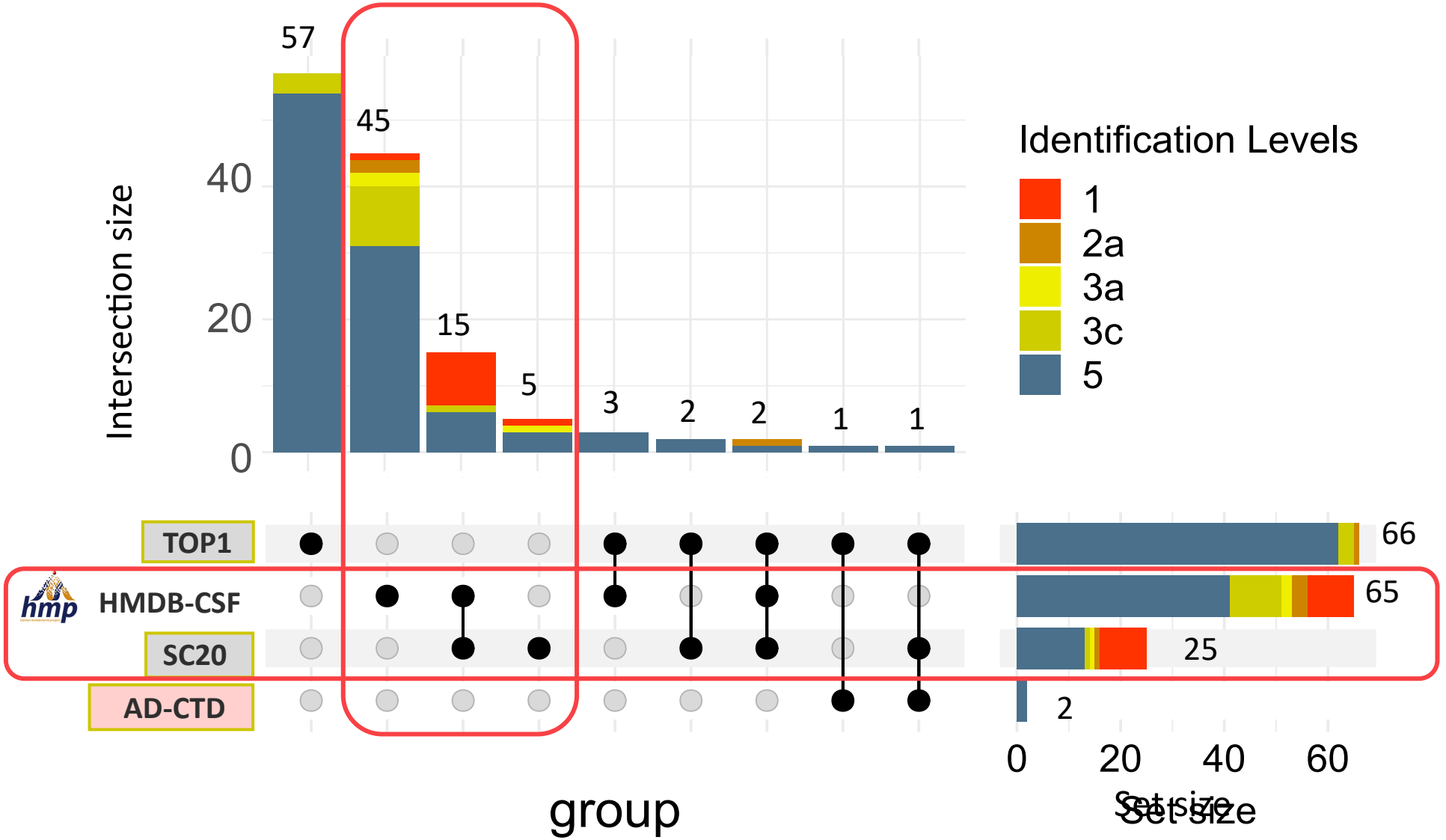
National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, United States



Kim S. et al. (2022). DOI: 10.1016/j.jmb.2022.167514 & Zaslavsky *et al.* (2022) DOI: 10.3389/frma.2021.689059

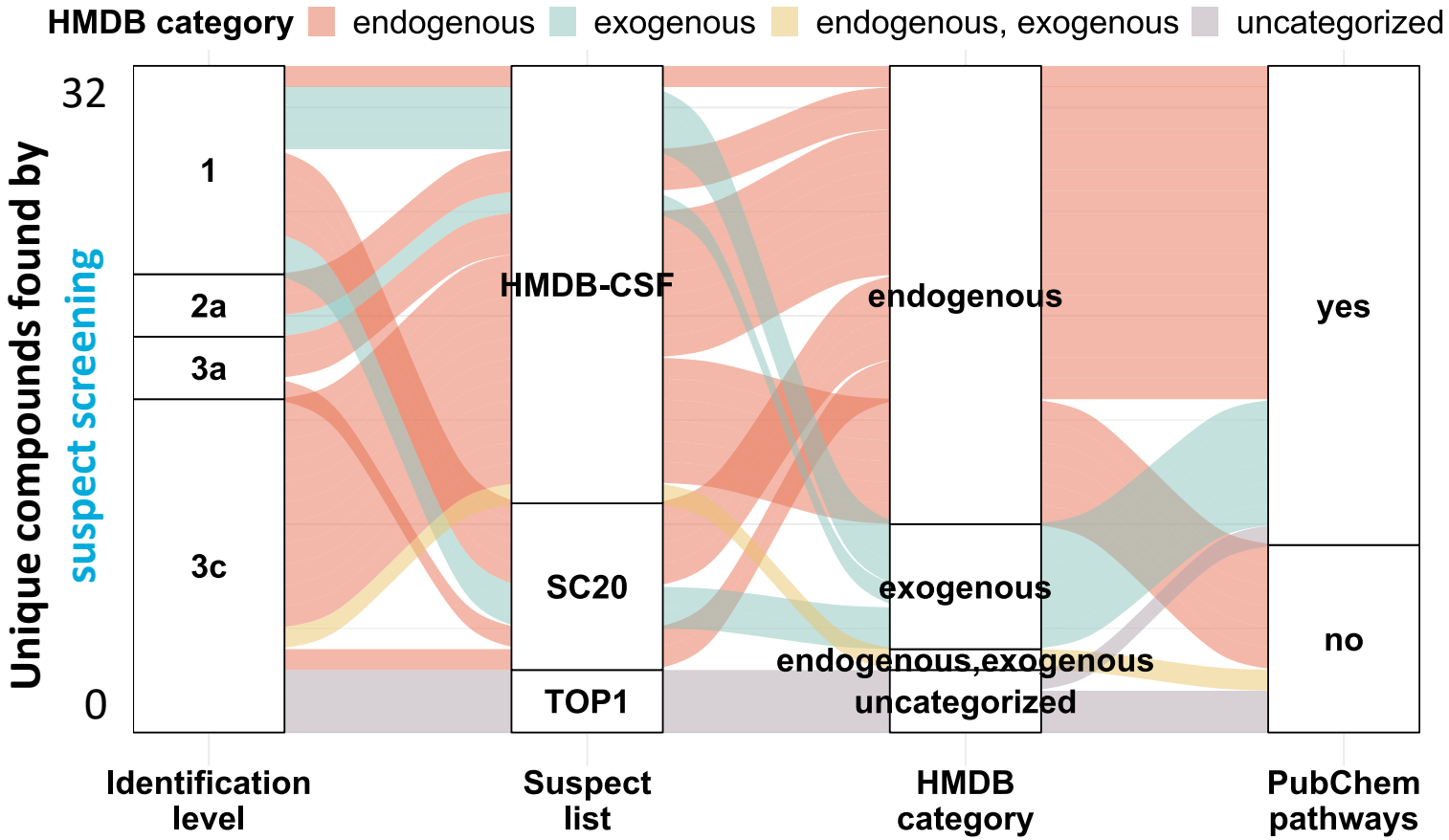
Complementarity of Suspect Lists





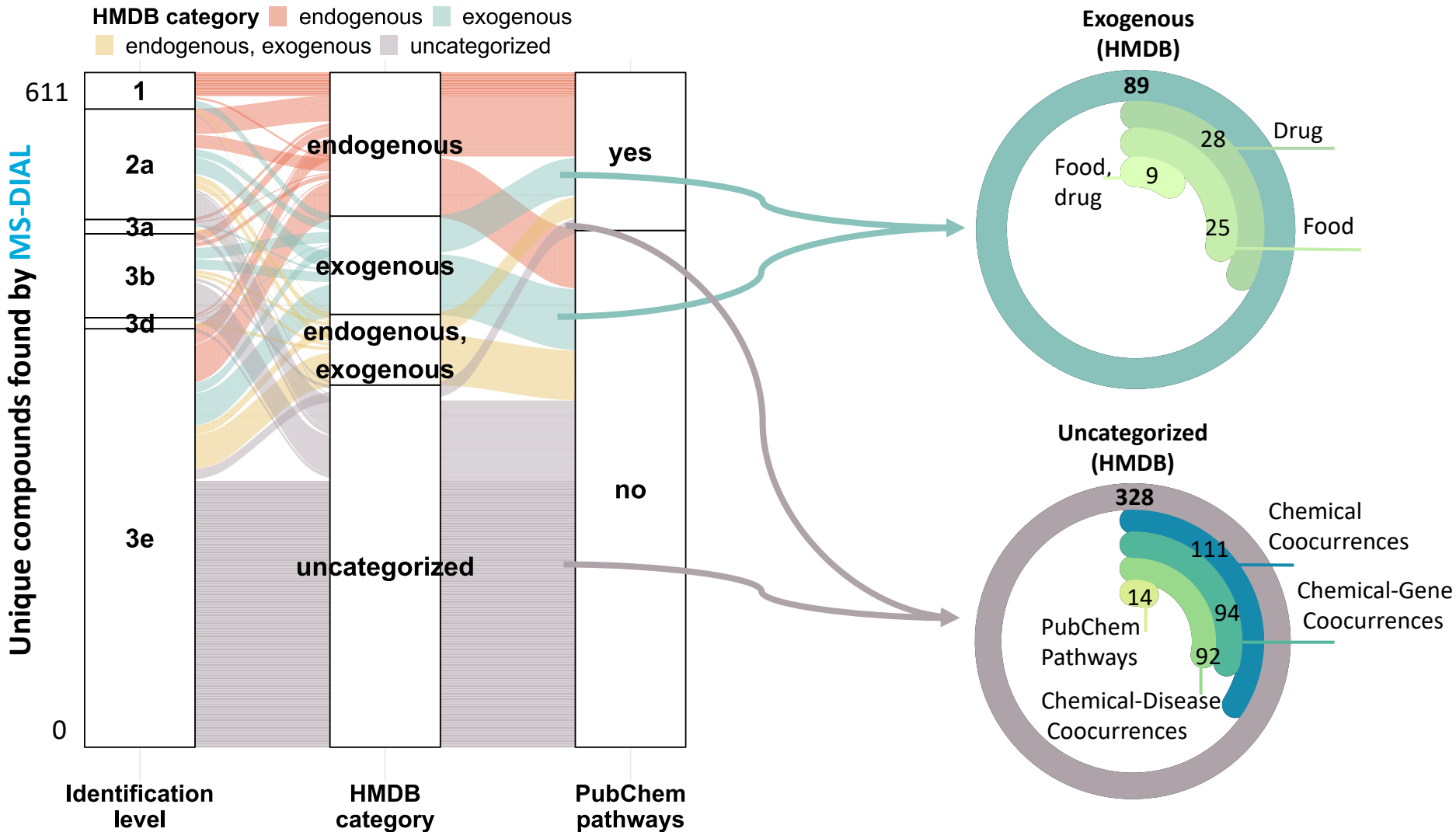
Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

Origins & Information about Annotations – Suspect Screening



Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

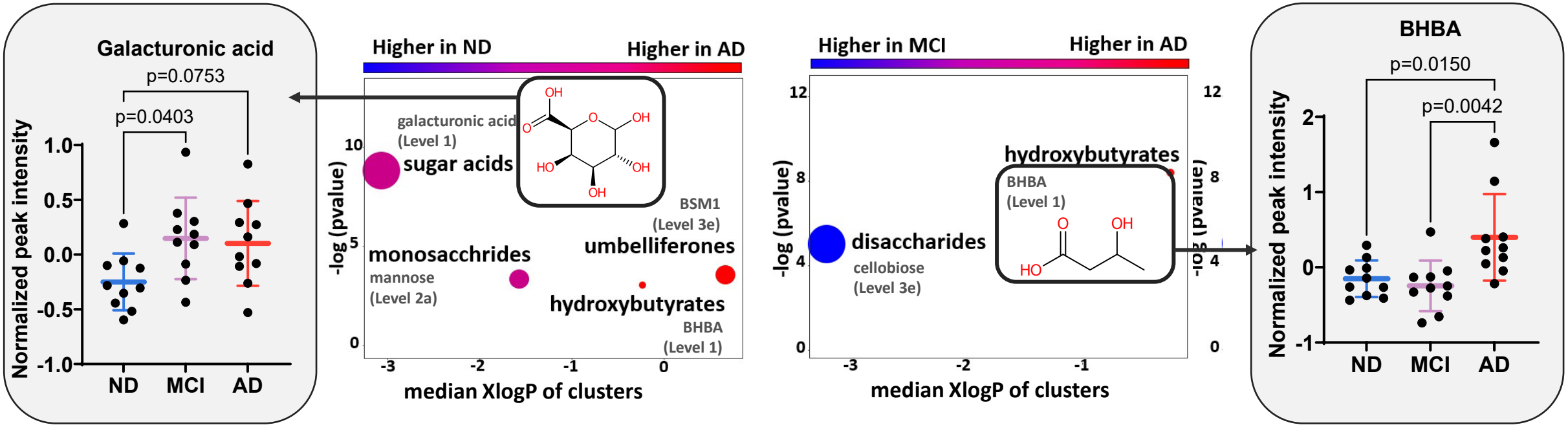
Origins & Information about Annotations – MS-DIAL MS/MS



Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

Chemical Similarity Enrichment Analysis (ChemRICH)

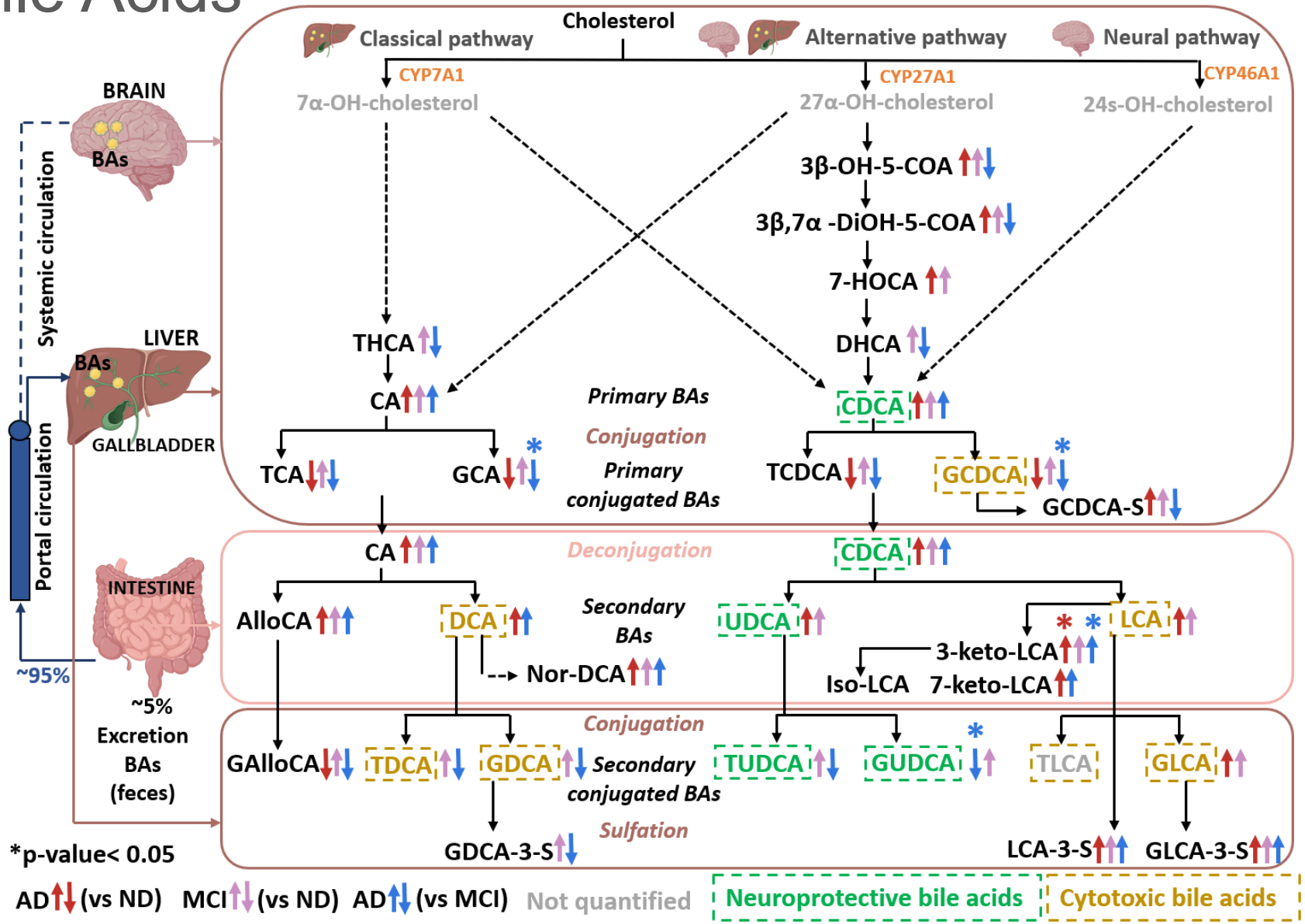
MICROH



BHBA: 3-hydroxybutanoic acid

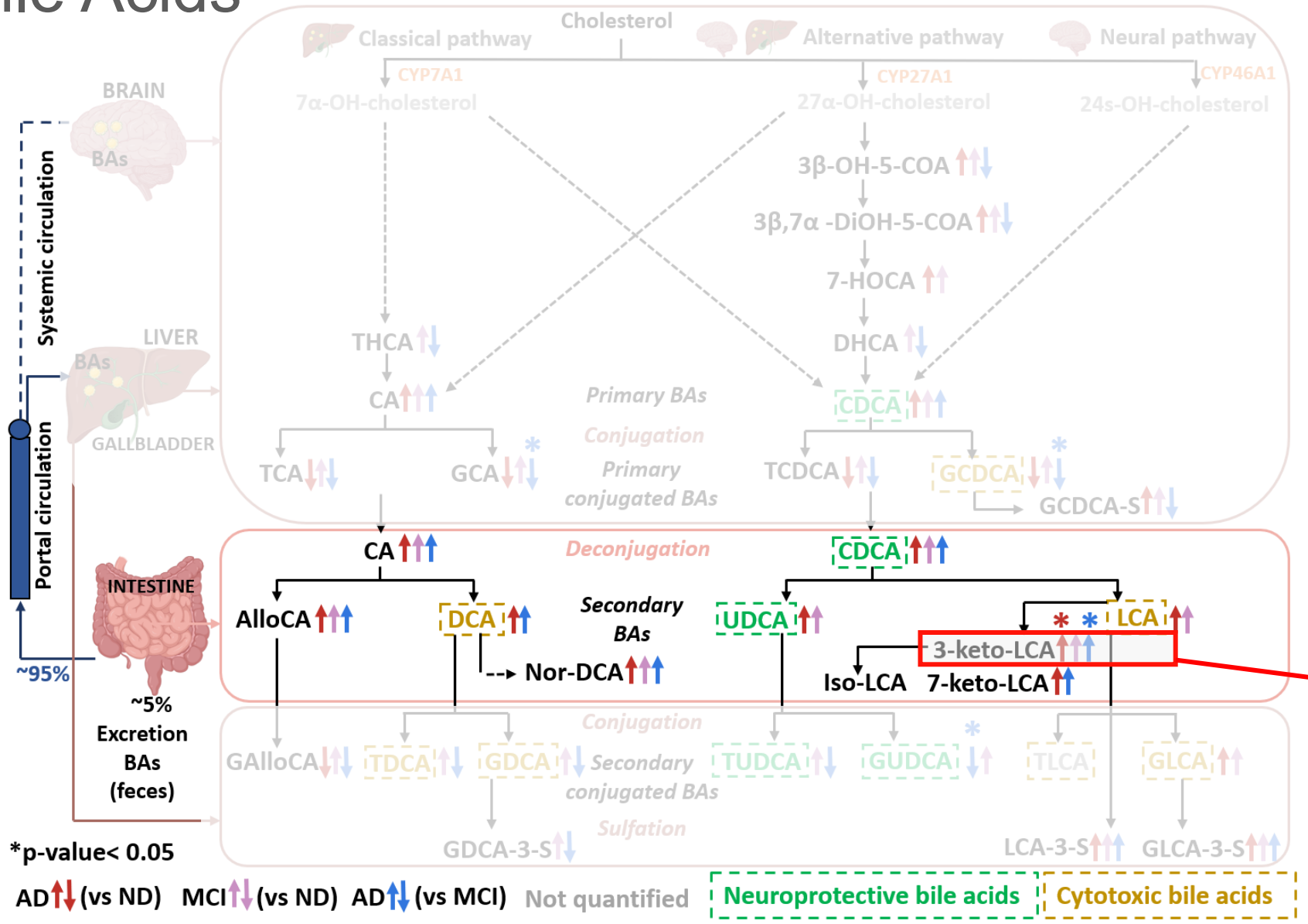
Talavera Andújar *et al*. (2024) ES&T, DOI: 10.1021/acs.est.3c10490
Barupal & Fiehn (2017) Sci. Rep. DOI: 10.1038/s41598-017-15231-w & https://chemrich.idsl.me/

Bile Acids

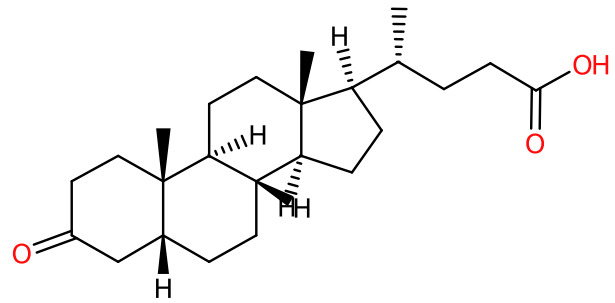


Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

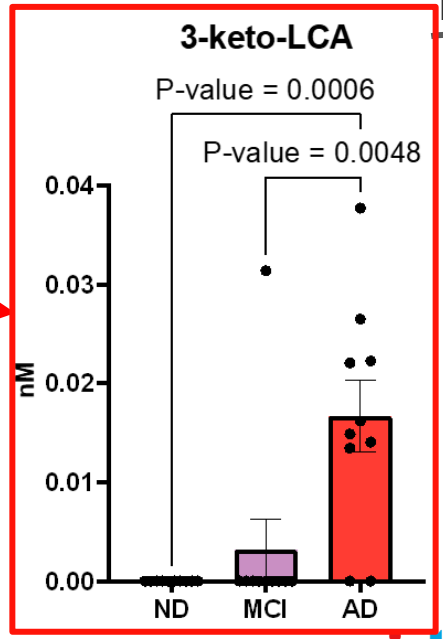
Bile Acids



3-keto-LCA: dehydrolithocholic acid



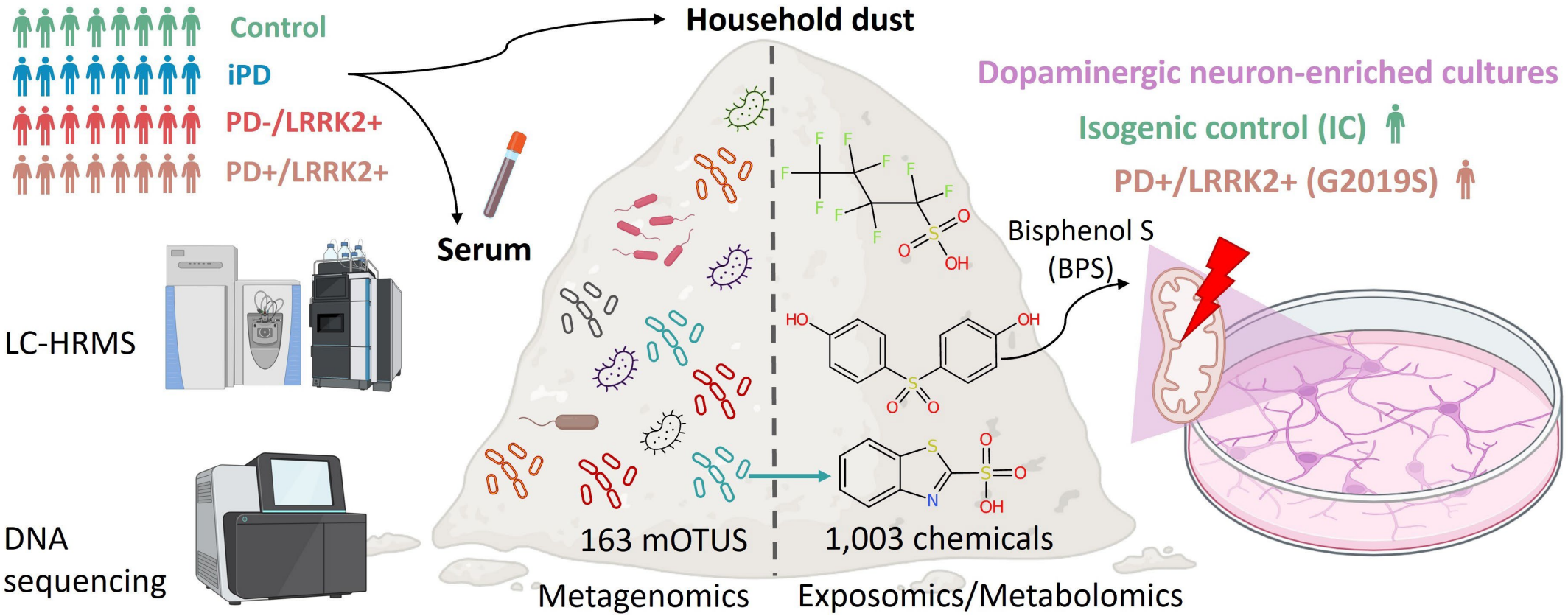
not yet



HMDB-CSF

Talavera Andújar *et al.* (2024) ES&T, DOI: 10.1021/acs.est.3c10490

Exposomics Case Study (2): PD, LRRK2 Penetrance & Dust

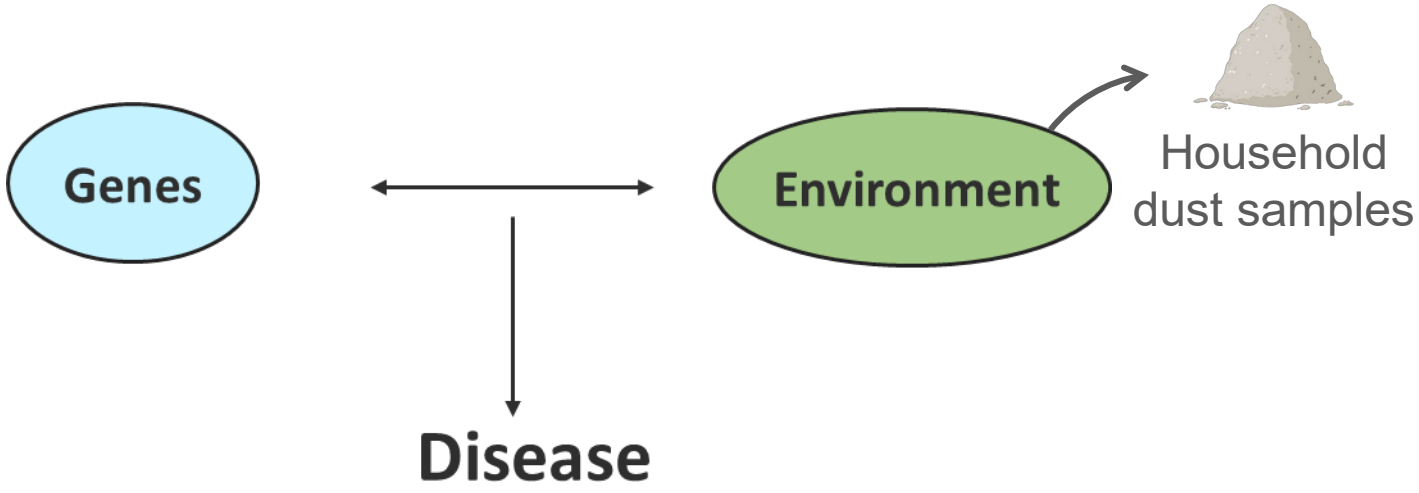


Talavera Andújar *et al*. (under review). Preprint: ChemRxiv DOI: 10.26434/chemrxiv-2024-rczsl

LRRK2-Associated Parkinson's Disease (PD)

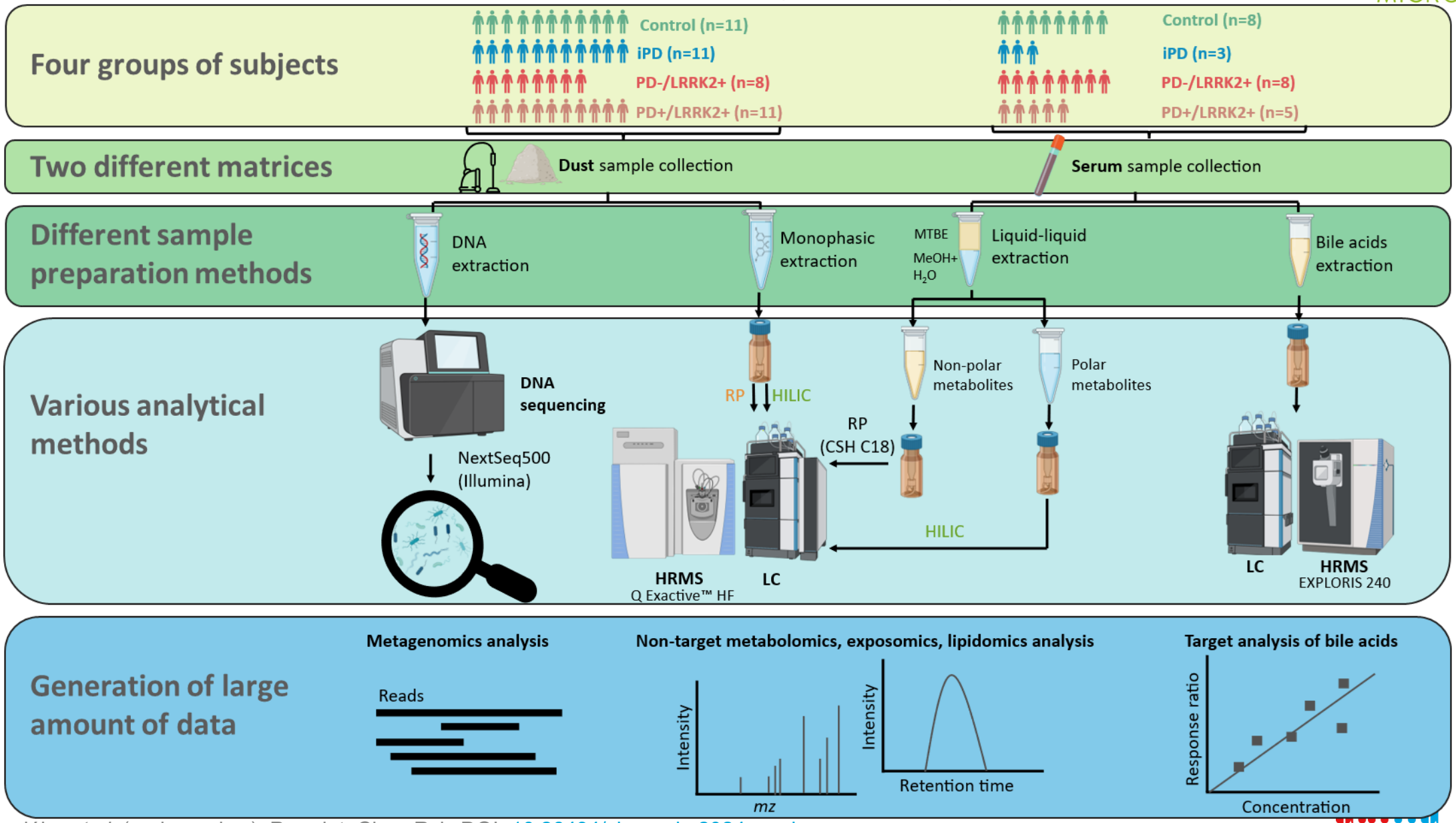
- Mutations in LRRK2 are the most common monogenetic cause of PD
  - 1-2 % of PD cases
- Penetrance of LRRK2 is age dependent and differs across geographic regions
  - "Reduced Penetrance"

LRRK2 p.G2019S carriers, **age of 60**



Source: Begoña Talavera Andújar. Refs: Hentati et al. (2014). DOI:10.1212/WNL.0000000000000675, Usnich et al. (2021). DOI: 10.3389/fneur.2021.710572 & Trinh J. et al. (2022). DOI:10.1515/medgen-2022-2127

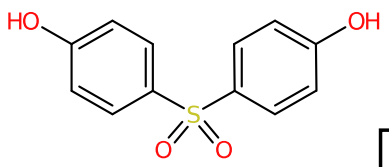
LRRK2, PD & Penetrance – Study Design

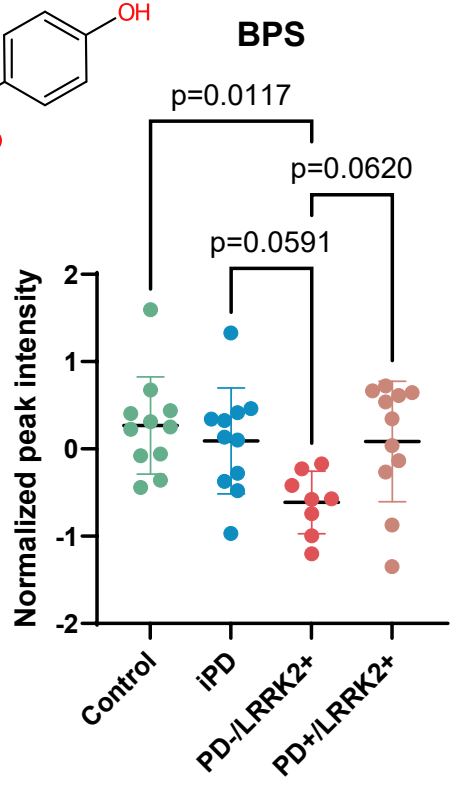


Talavera Andújar *et al*. (under review). Preprint: ChemRxiv DOI: 10.26434/chemrxiv-2024-rczsl

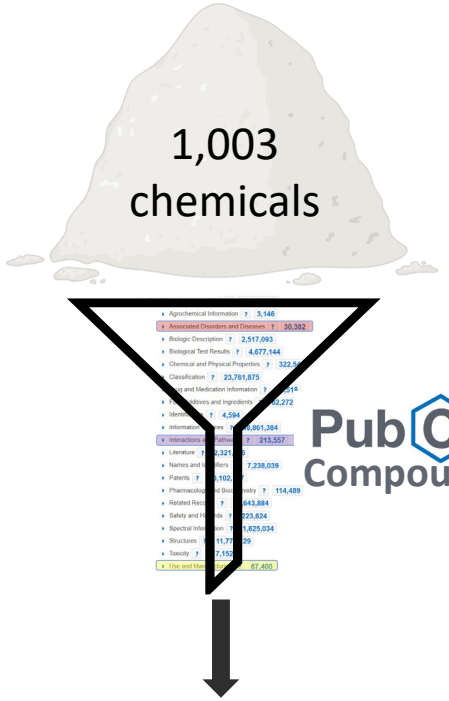
LRRK2 & PD – Hazardous Chemical Replacements in Dust

Bisphenol S





Emerged as a “safer” alternative to bisphenol A (BPA)



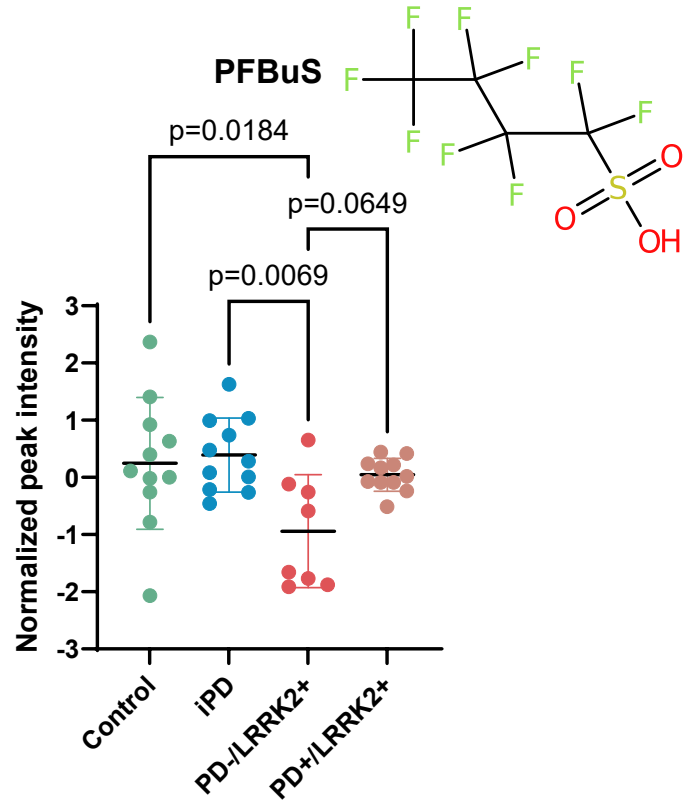
Use and manufacturing (549)

Associated Disease/Disorder (543)

Pathways (331)

Hazardous chemical replacements (2)

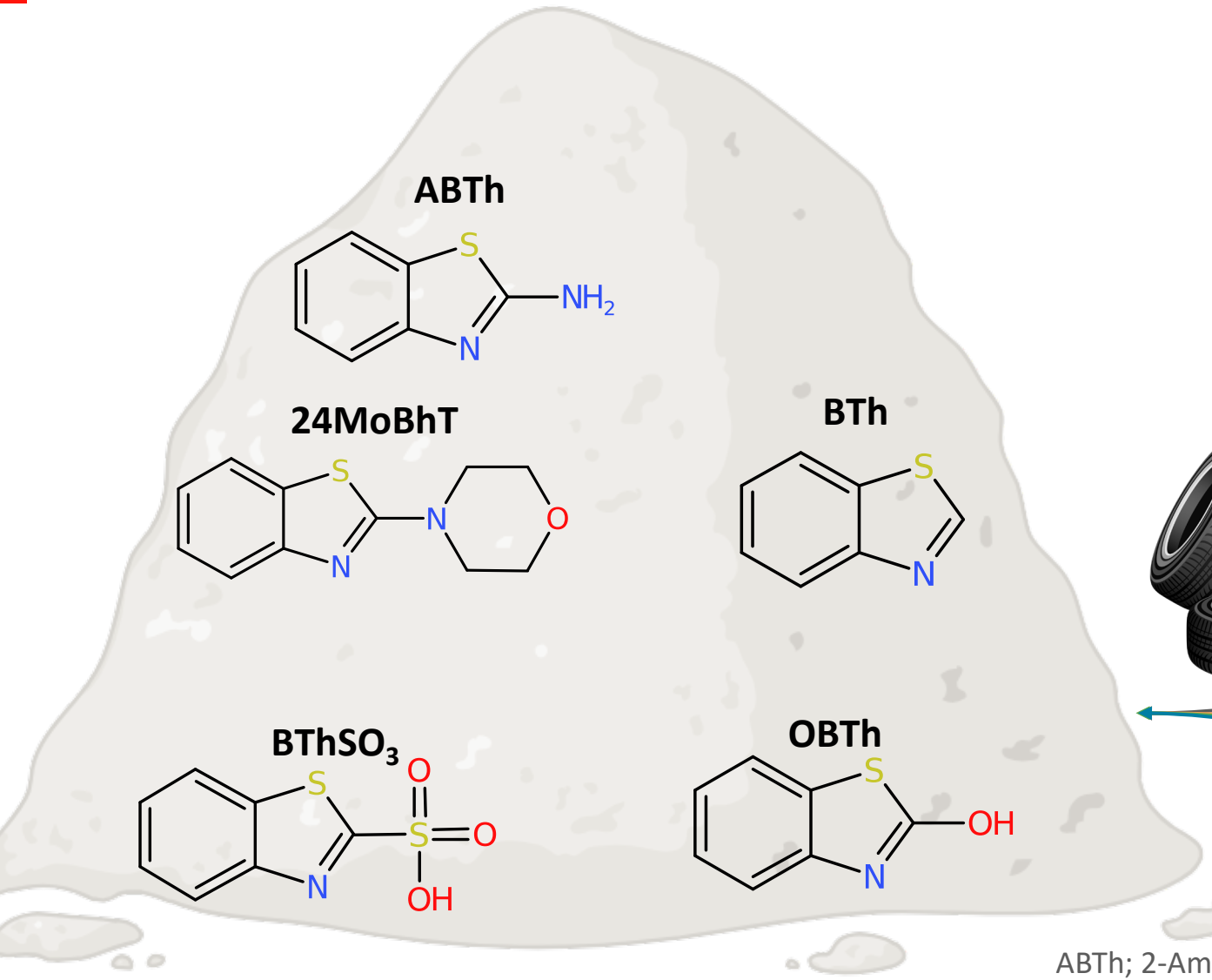
Perfluorobutane Sulfonic Acid

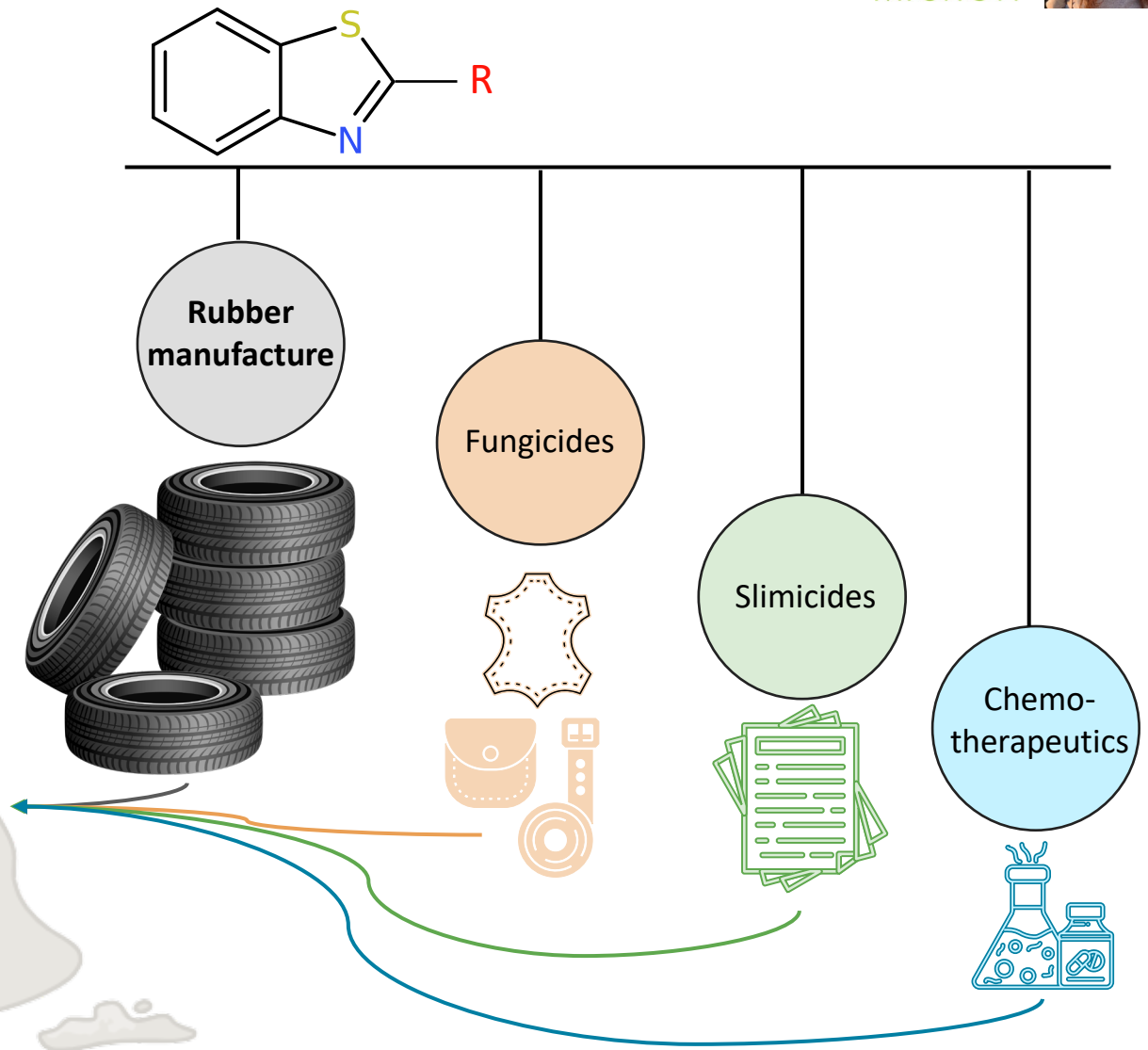


PFBuS is a substitute for perfluorooctanesulfonate (PFOS)

Talavera Andújar *et al*. (under review). Preprint: ChemRxiv DOI: 10.26434/chemrxiv-2024-rczsl

LRRK2 & PD – Surprising Number of Benzothiazoles

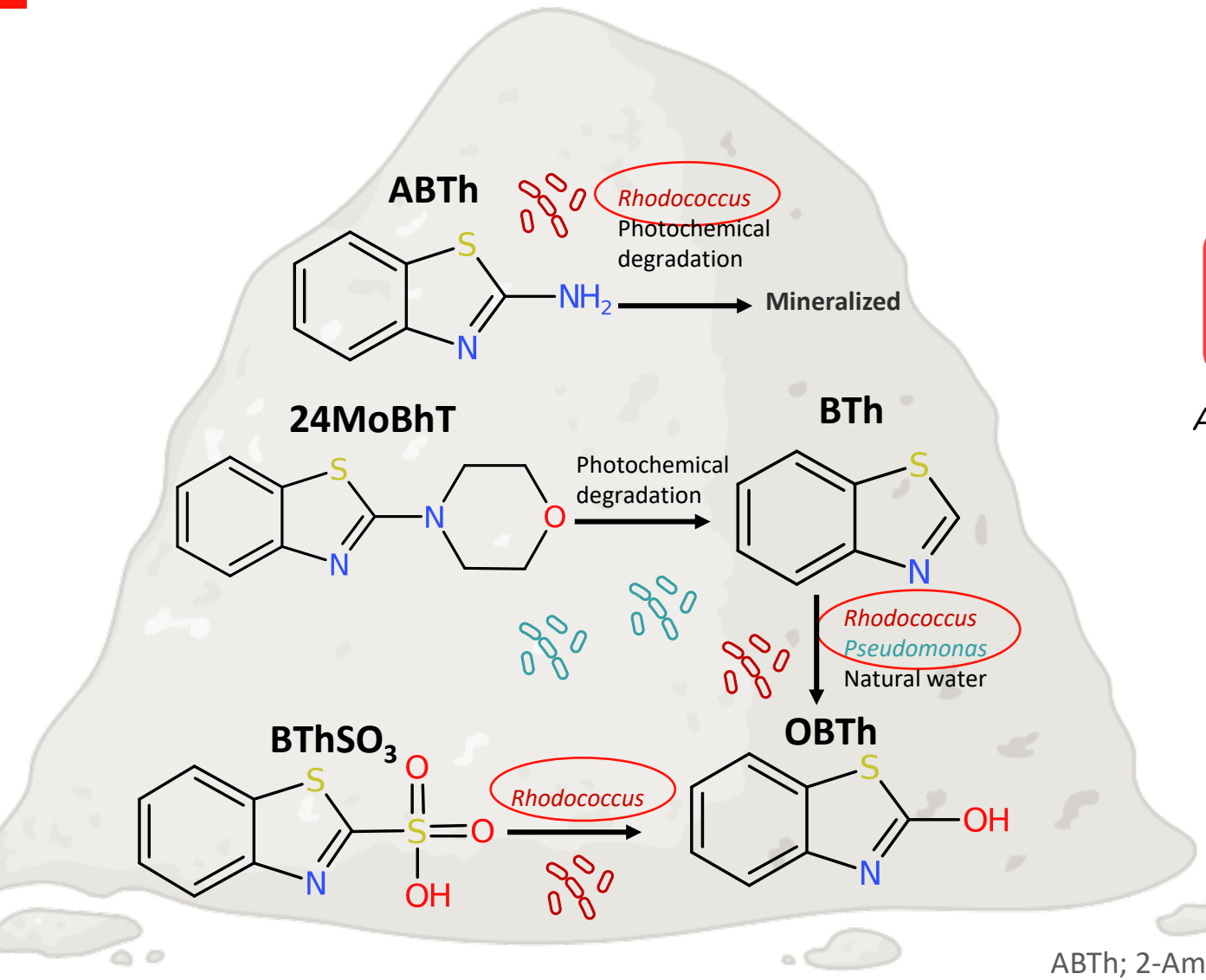


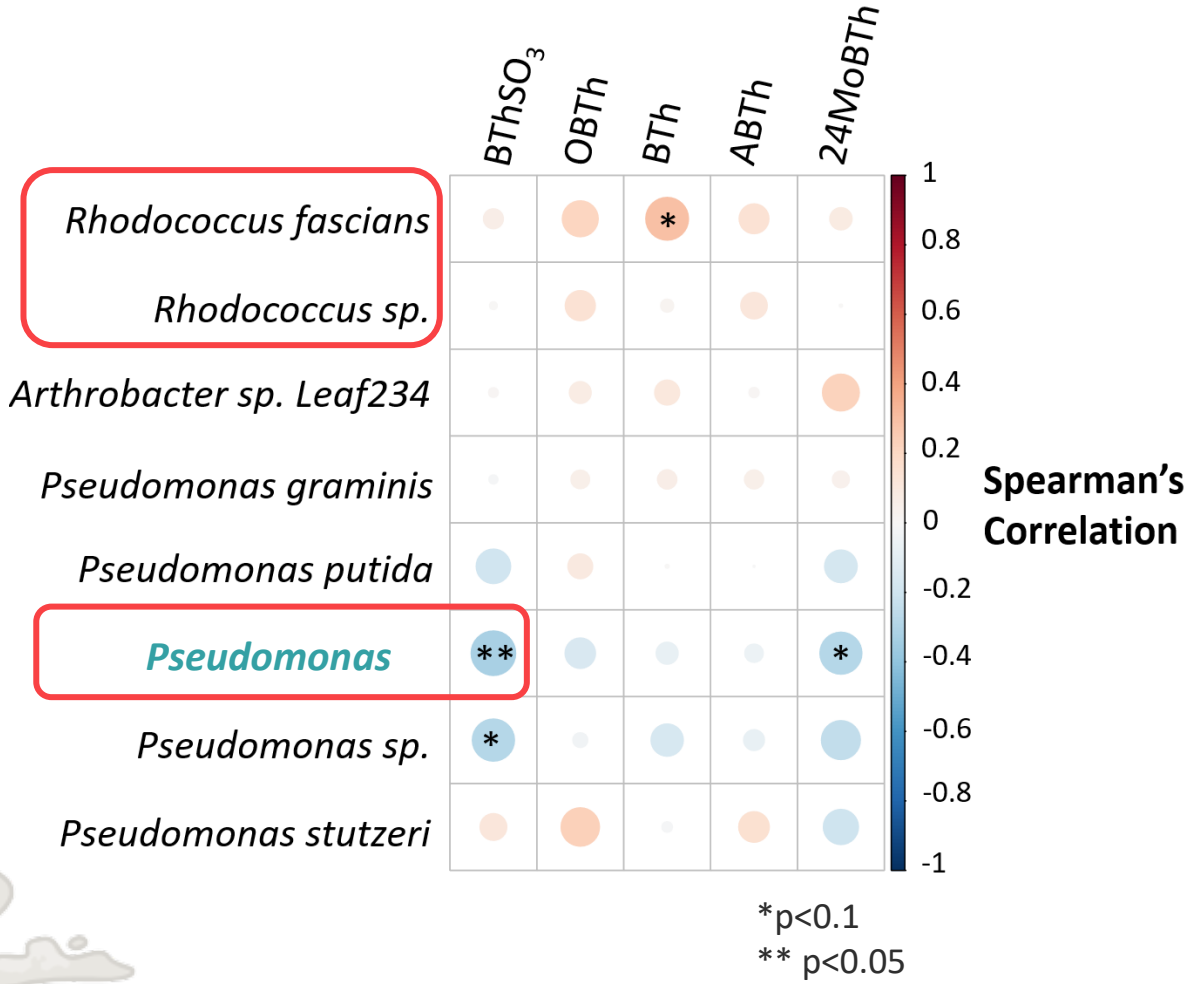


ABTh; 2-Aminobenzothiazole, BTh; benzothiazole, $BThSO_3$; 2-Benzothiazolesulfonic acid, OBTh; 2-hydroxybenzothiazole, and 24MoBTh; 2-(4-Morpholinyl)benzothiazole

Talavera Andújar *et al*. (under review). Preprint: ChemRxiv DOI: 10.26434/chemrxiv-2024-rczsl

LRRK2 & PD – Benzothiazoles and Microbes?





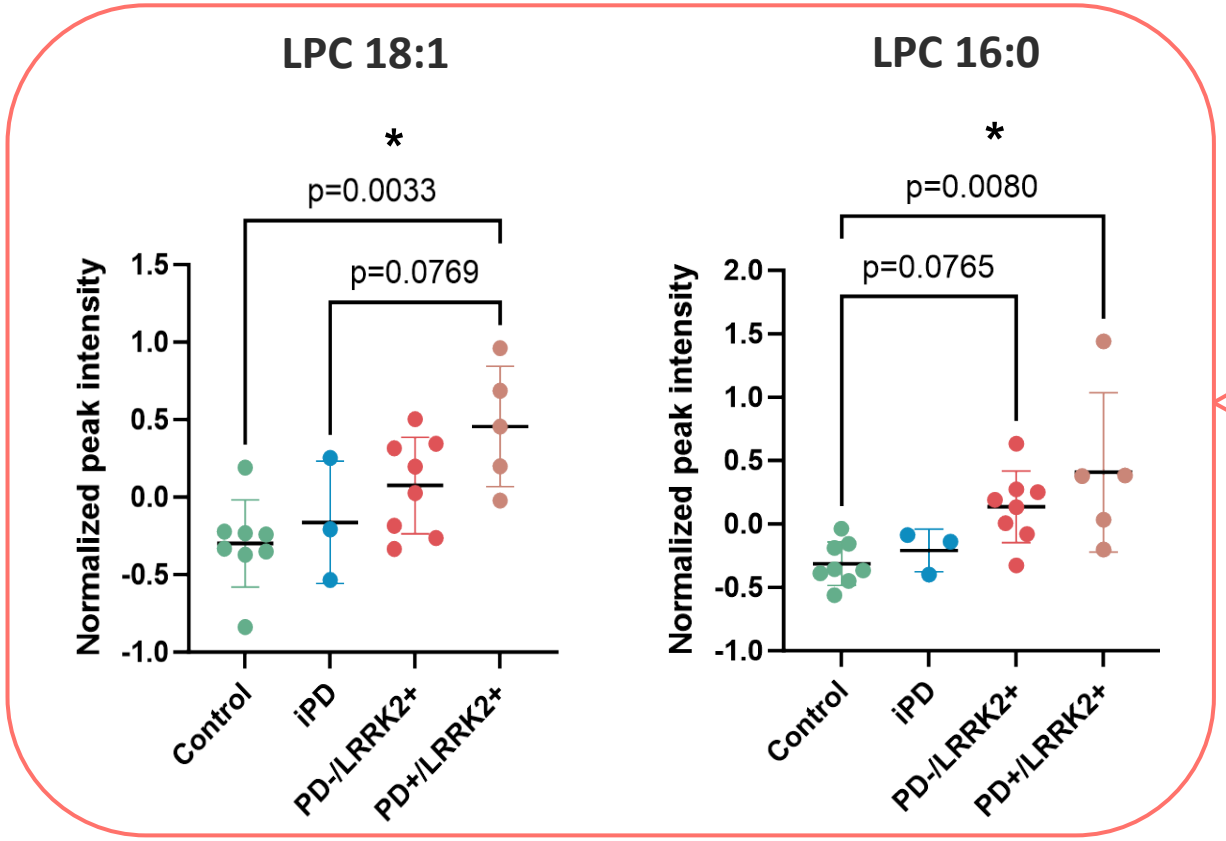
ABTh; 2-Aminobenzothiazole, BTh; benzothiazole, $BThSO_3$; 2-Benzothiazolesulfonic acid, OBTh; 2-hydroxybenzothiazole, and 24MoBTh; 2-(4-Morpholinyl)benzothiazole

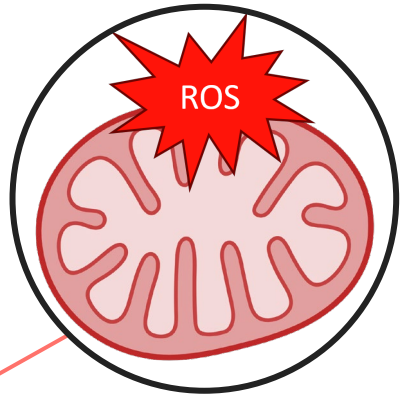
Talavera Andújar *et al*. (under review). Preprint: ChemRxiv DOI: 10.26434/chemrxiv-2024-rczsl

LRRK2 & PD – Lipid Alterations in Serum – BPS Exposure?

Total lipids (Level 2-3) — 313

Statistically significant — 43





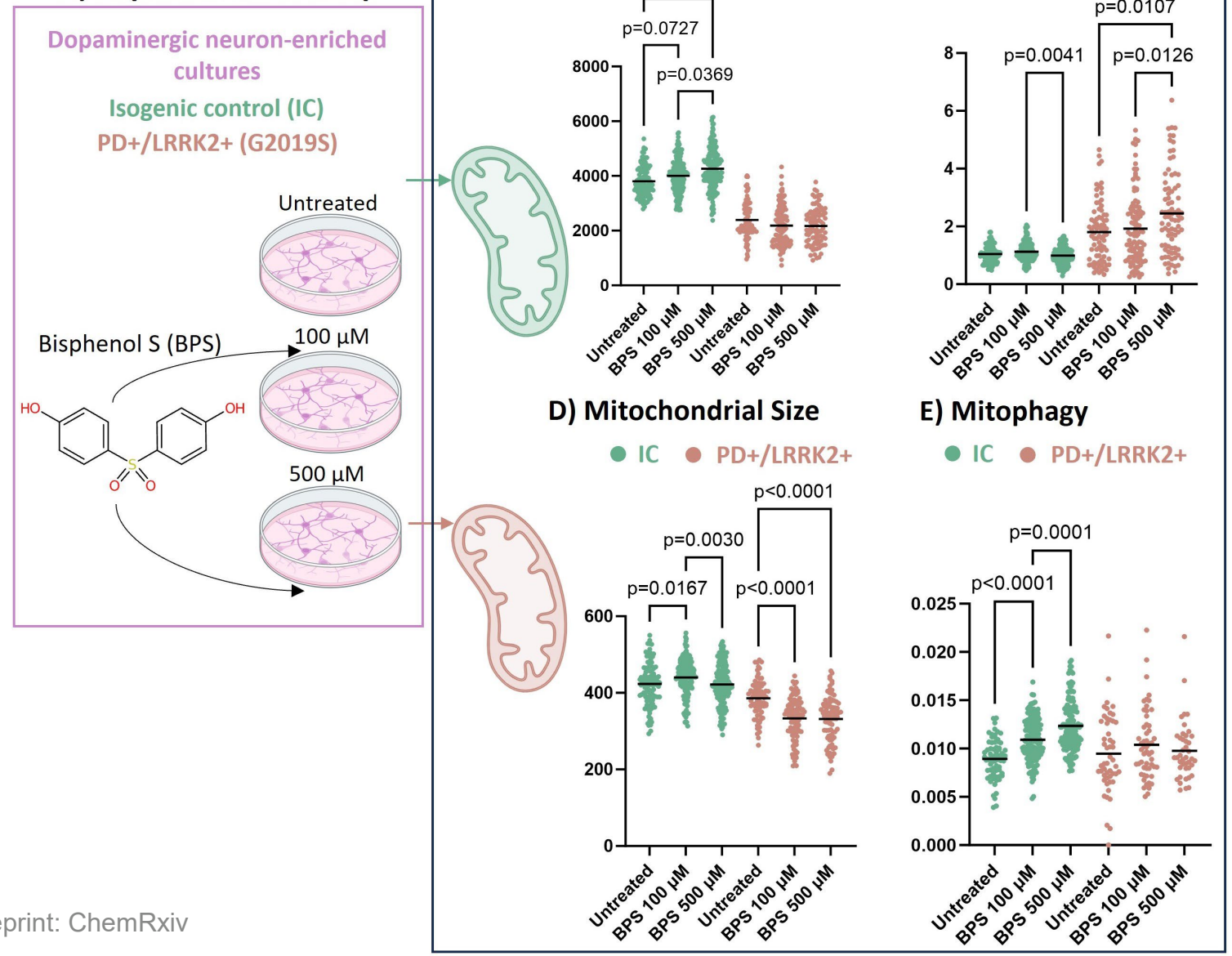
Oxidative stress and mitochondrial dysfunction

Neuroinflammation

Talavera Andújar *et al*. (under review). Preprint: ChemRxiv DOI: 10.26434/chemrxiv-2024-rczsl

LPC: Lysophosphatidylcholine

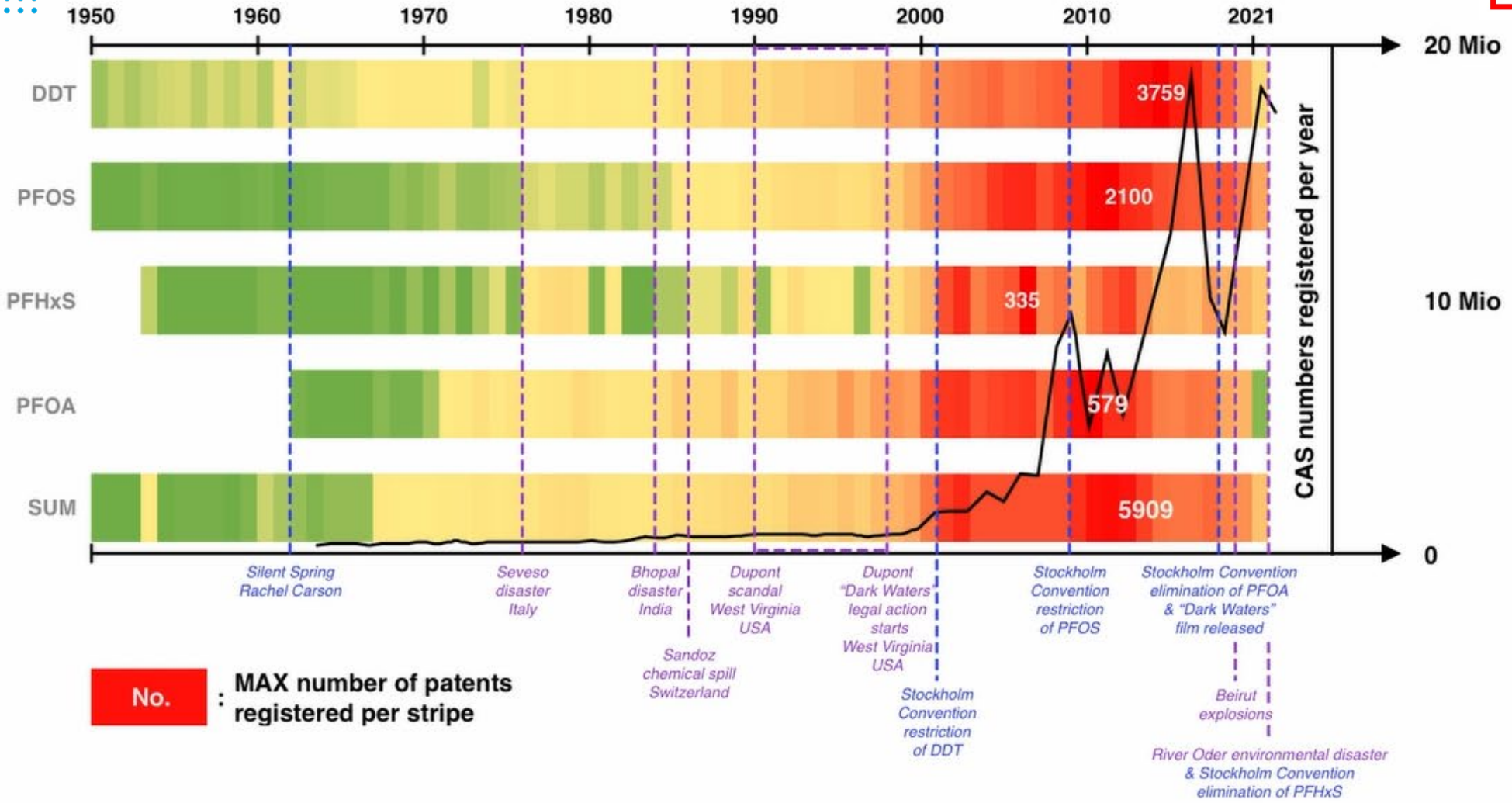
BPS Exposure



Talavera Andújar *et al*. (under review). Preprint: ChemRxiv
DOI: 10.26434/chemrxiv-2024-rczsl

Our Chemical Past, Present, and Future

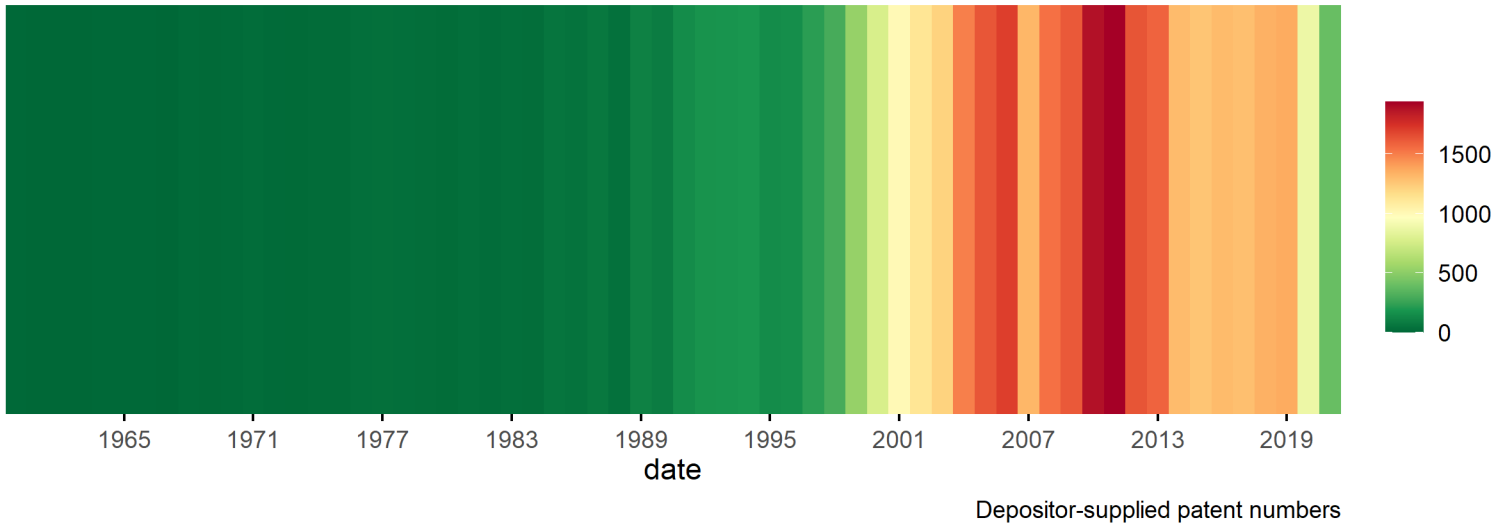
low / medium / high number of patents registered per year (WIPO)

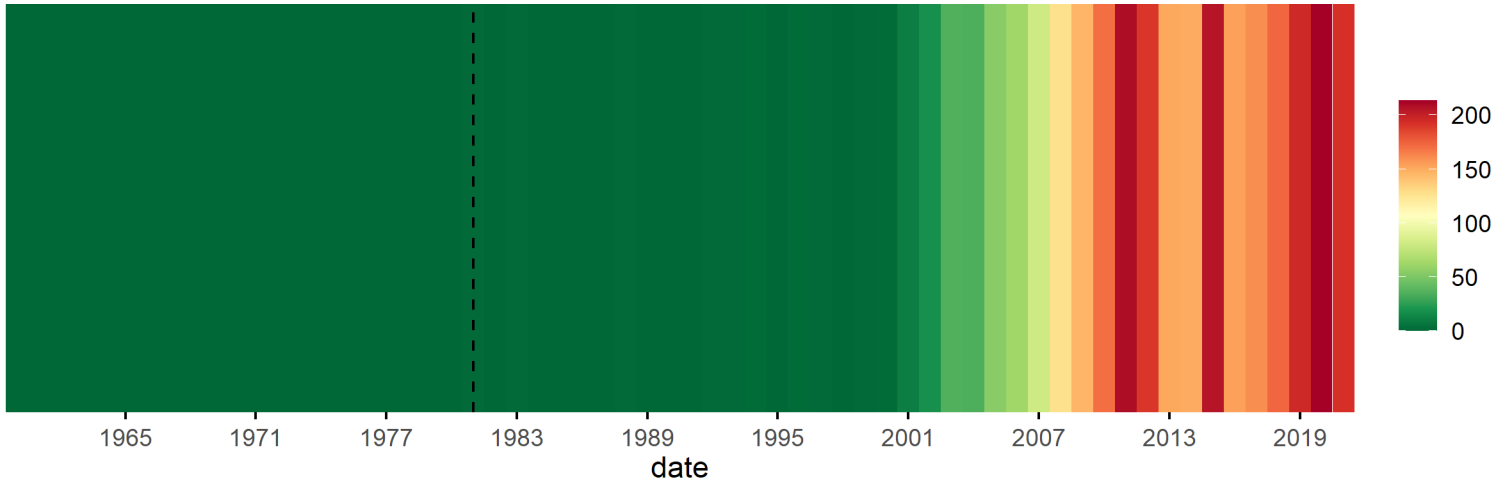


Composition & Animation: Jamie Perera; Figure: Dagny Aurich; mod. from Arp et al. (2023) DOI: 10.1021/acs.est.3c01735. Credits: Jamie Perera, Dagny Aurich, Paul Thiessen, Emma Schymanski

Chemical Stripes

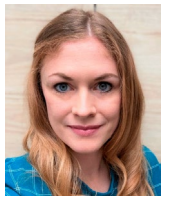
Patents & Literature





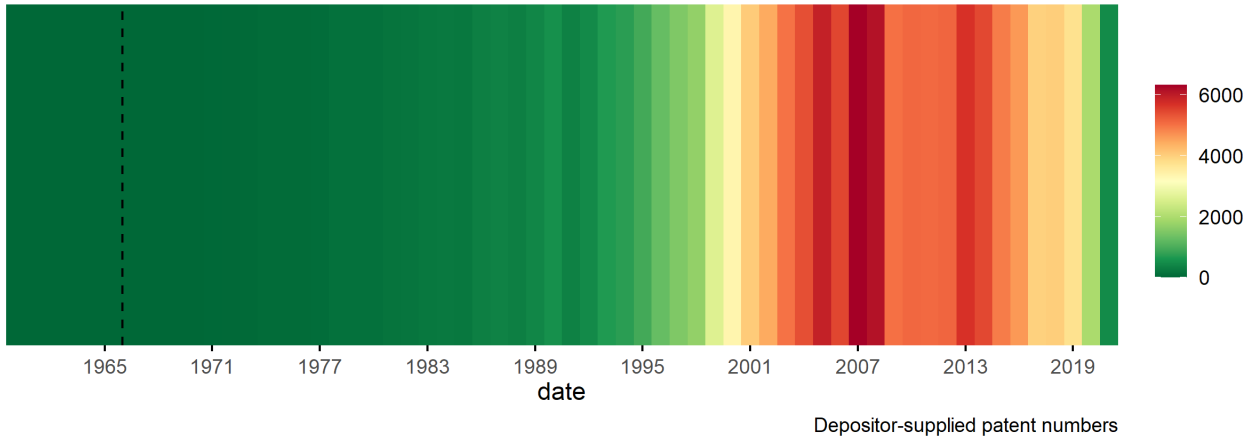
Aurich *et al.* (2024) ES&T Letters DOI: 10.1021/acs.estlett.4c00560
https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

Variety of Stripe Patterns: Patents



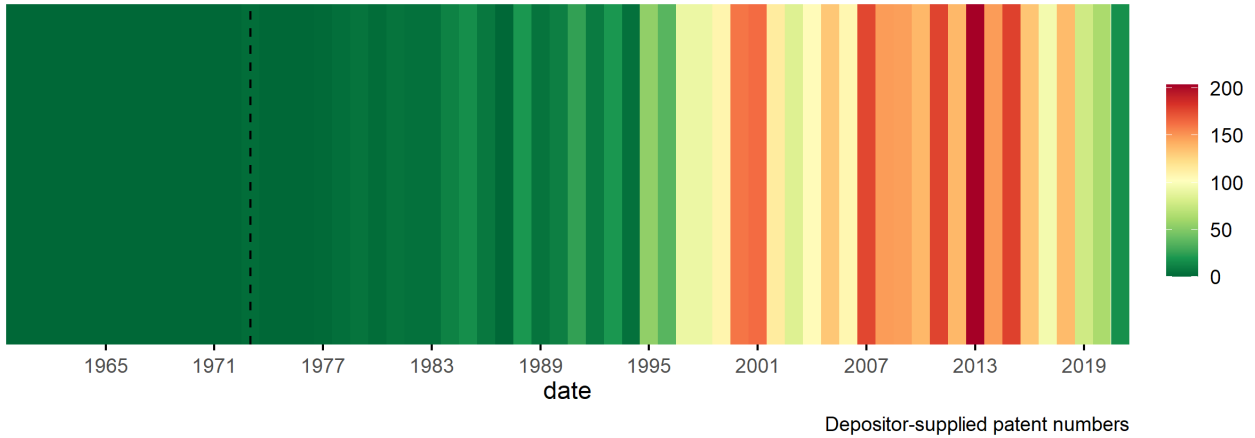
Chemical Stripes for Caffeine

PubChem CID: 2519
First patent: 1966



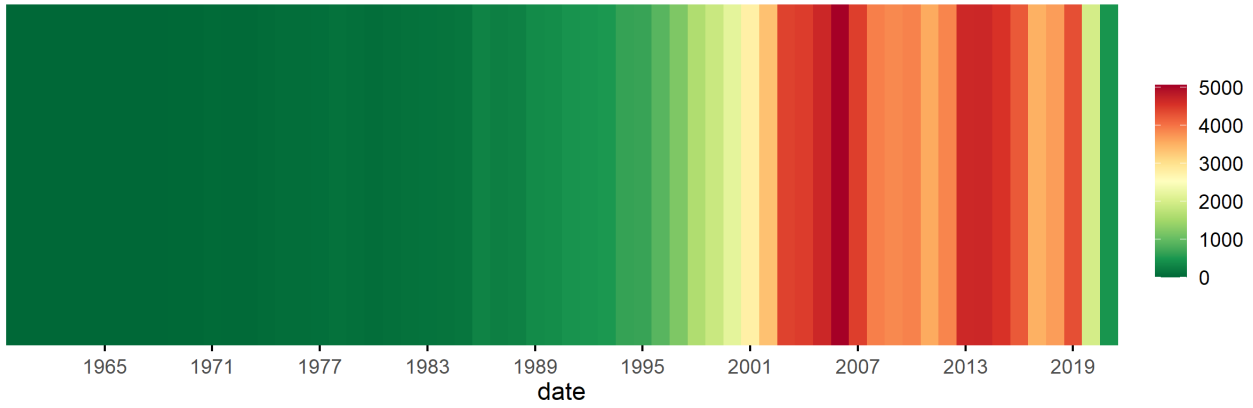
Chemical Stripes for Acetylcarnitine

PubChem CID: 1
First patent: 1973



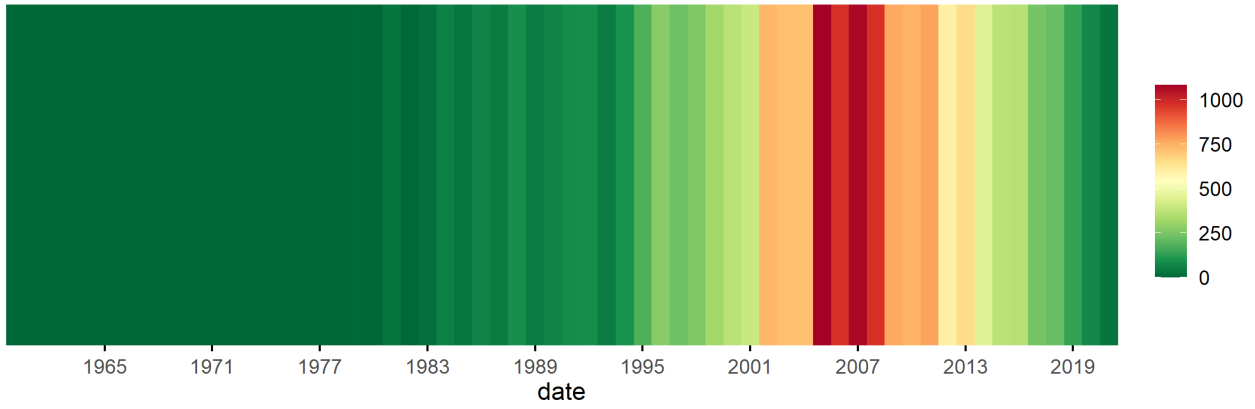
Chemical Stripes for Nicotine

PubChem CID: 89594
First patent: 1954



Chemical Stripes for Gallopamil

PubChem CID: 1234
First patent: 1951

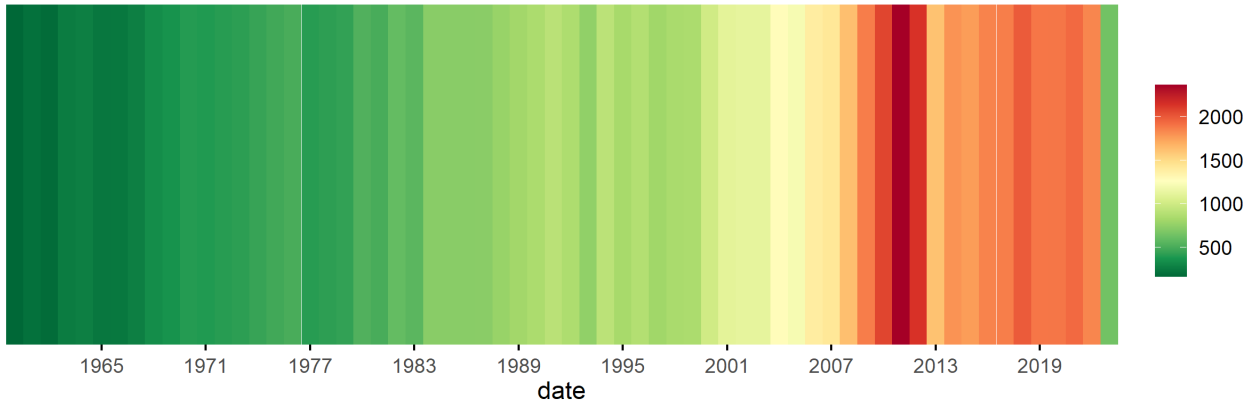


https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

Variety of Stripe Patterns: Literature

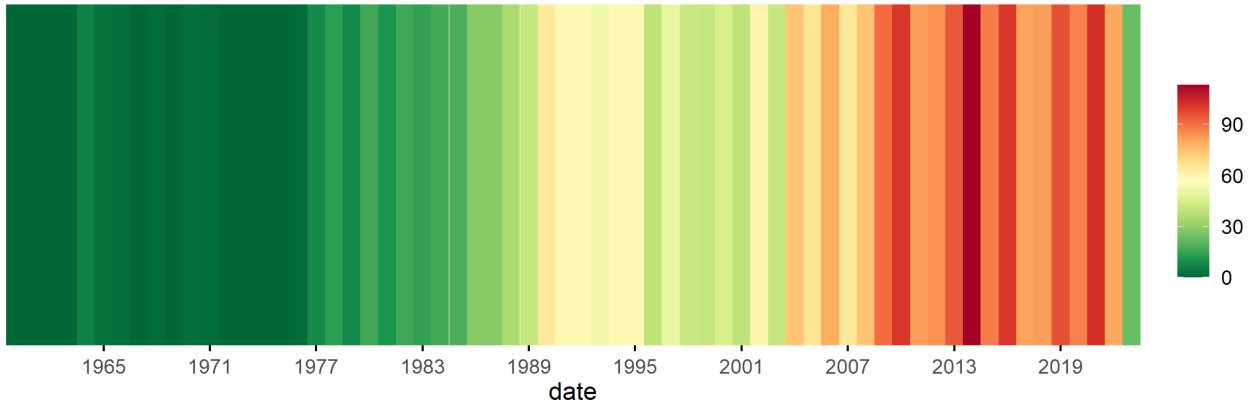
Chemical Stripes for Caffeine

PubChem CID: 2519
First reference: 1851



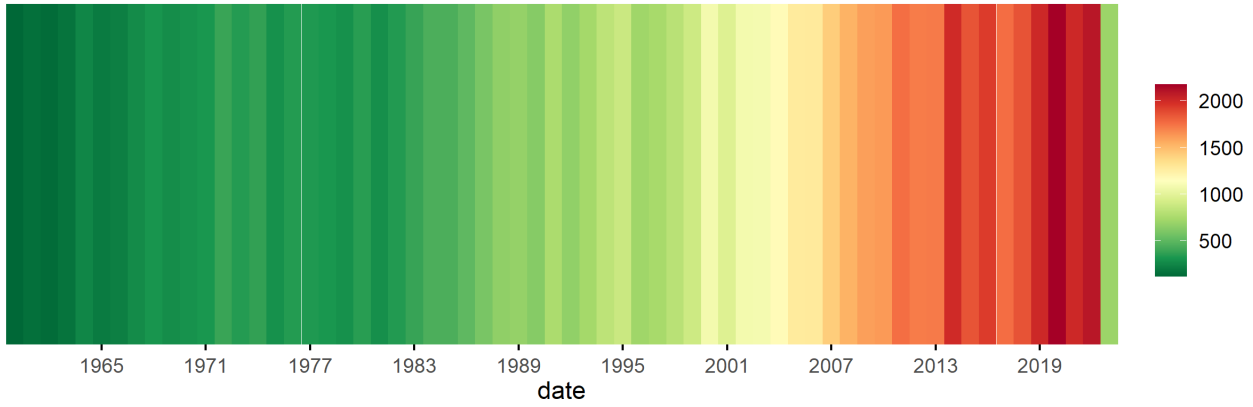
Chemical Stripes for Acetylcarnitine

PubChem CID: 1
First reference: 1959



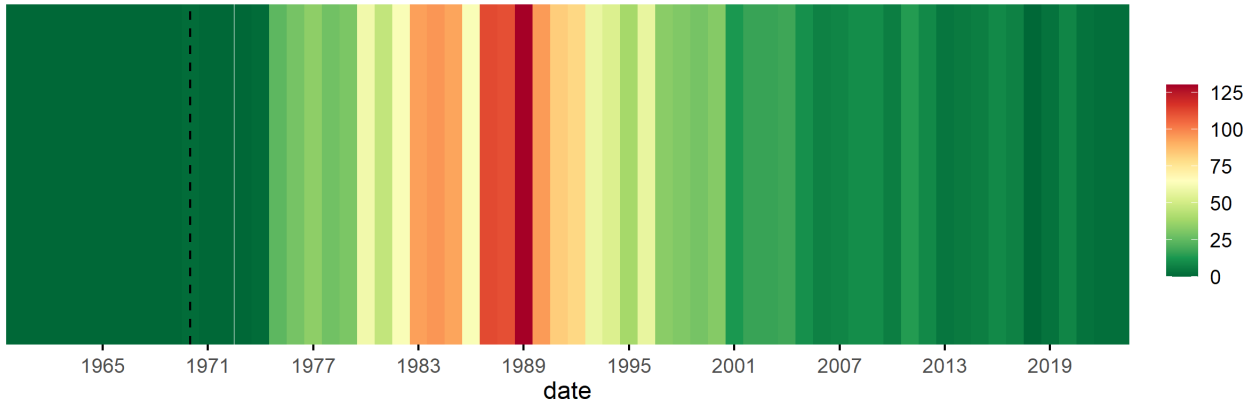
Chemical Stripes for Nicotine

PubChem CID: 89594
First reference: 1854



Chemical Stripes for Gallopamil

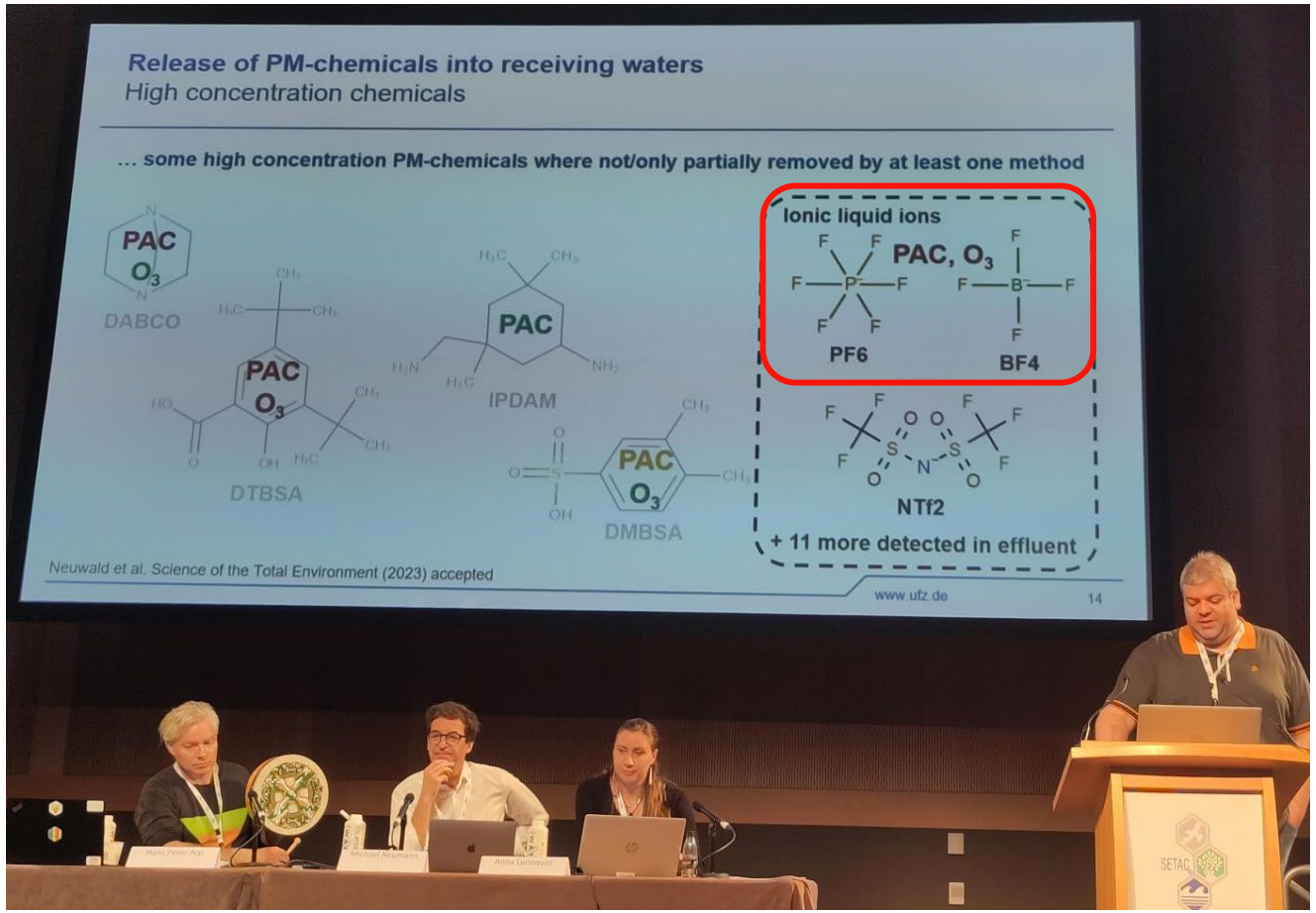
PubChem CID: 1234
First reference: 1970

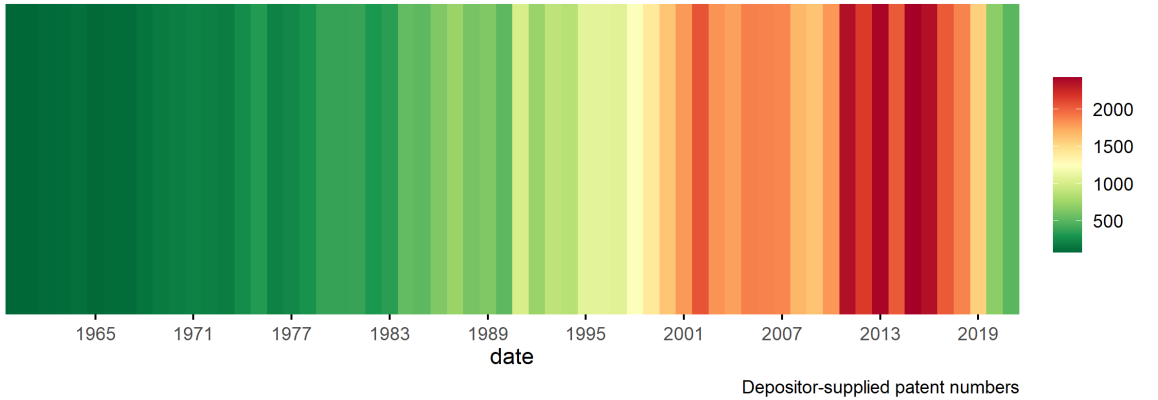


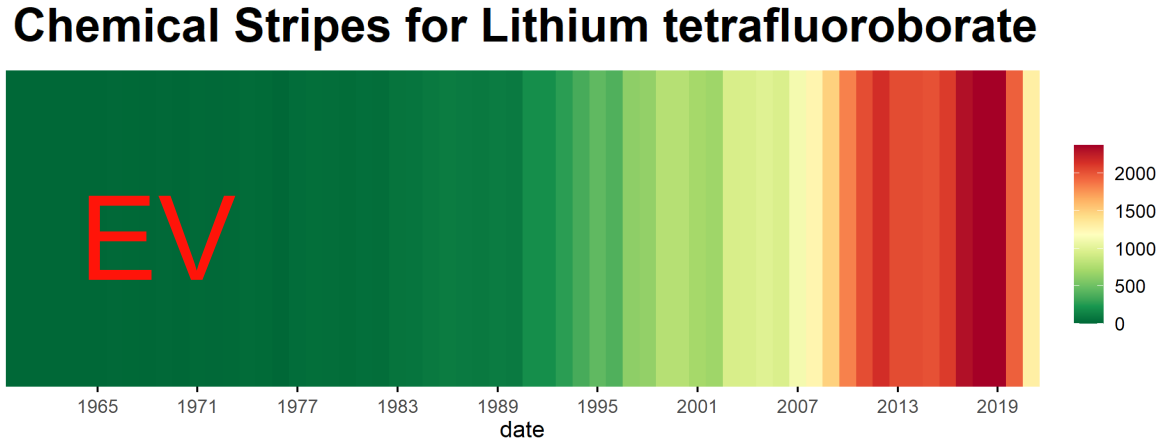
https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

The Chemical Stripes and Patent Data

(live stripe calculations during SETAC 2023)







Left: Neuwald *et al, STOTEN*, DOI: [10.1016/j.scitotenv.2023.163921](https://doi.org/10.1016/j.scitotenv.2023.163921)
Photo of Daniel Zahn, UFZ at SETAC Europe, 30 April – 4 May, 2023. Image reused with permission

Subsetting PubChem

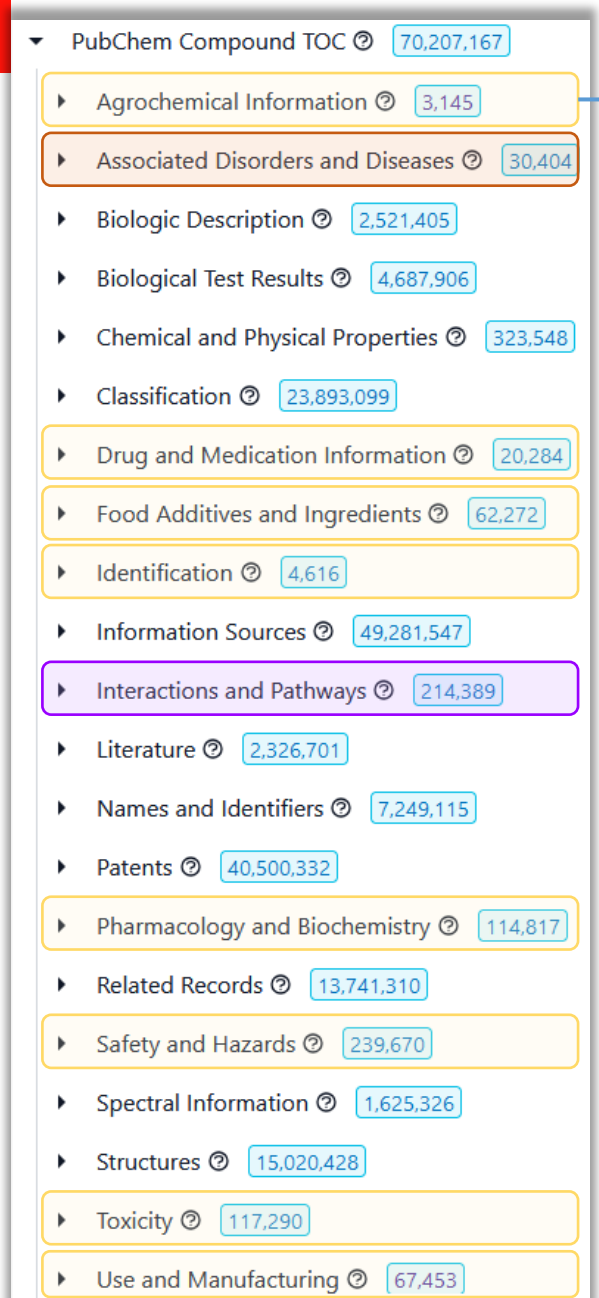
- PubChem Compound TOC 70,207,167
  - Agrochemical Information 3,145
  - Associated Disorders and Diseases 30,404
  - Biologic Description 2,521,405
  - Biological Test Results 4,687,906
  - Chemical and Physical Properties 323,548
  - Classification 23,893,099
  - Drug and Medication Information 20,284
  - Food Additives and Ingredients 62,272
  - Identification 4,616
  - Information Sources 49,281,547
  - Interactions and Pathways 214,389
  - Literature 2,326,701
  - Names and Identifiers 7,249,115
  - Patents 40,500,332
  - Pharmacology and Biochemistry 114,817
  - Related Records 13,741,310
  - Safety and Hazards 239,670
  - Spectral Information 1,625,326
  - Structures 15,020,428
  - Toxicity 117,290
  - Use and Manufacturing 67,453

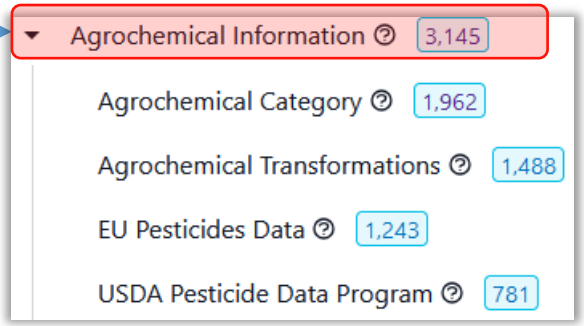
→

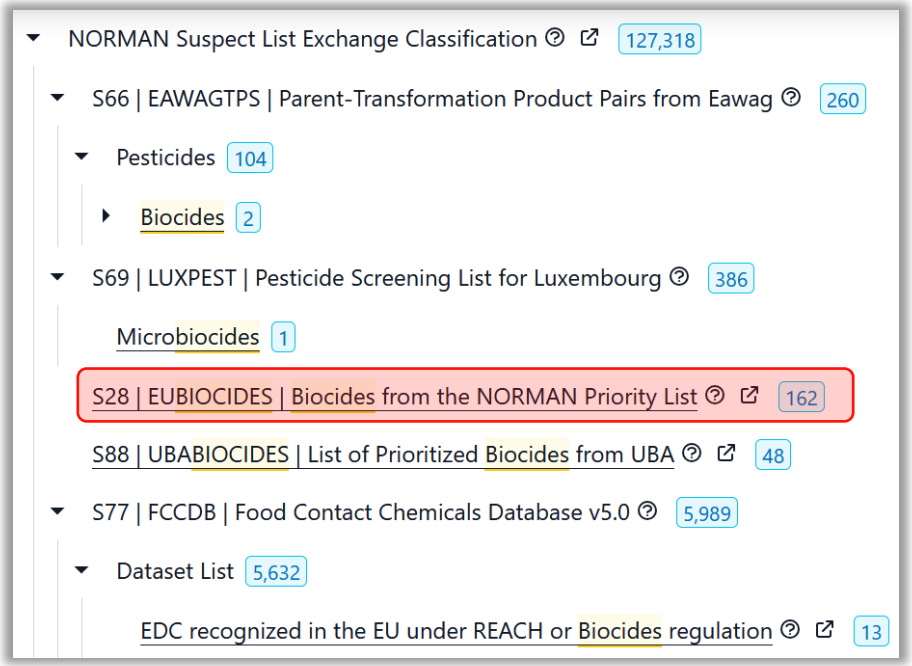
- Agrochemical Information 3,145
  - Agrochemical Category 1,962
  - Agrochemical Transformations 1,488
  - EU Pesticides Data 1,243
  - USDA Pesticide Data Program 781

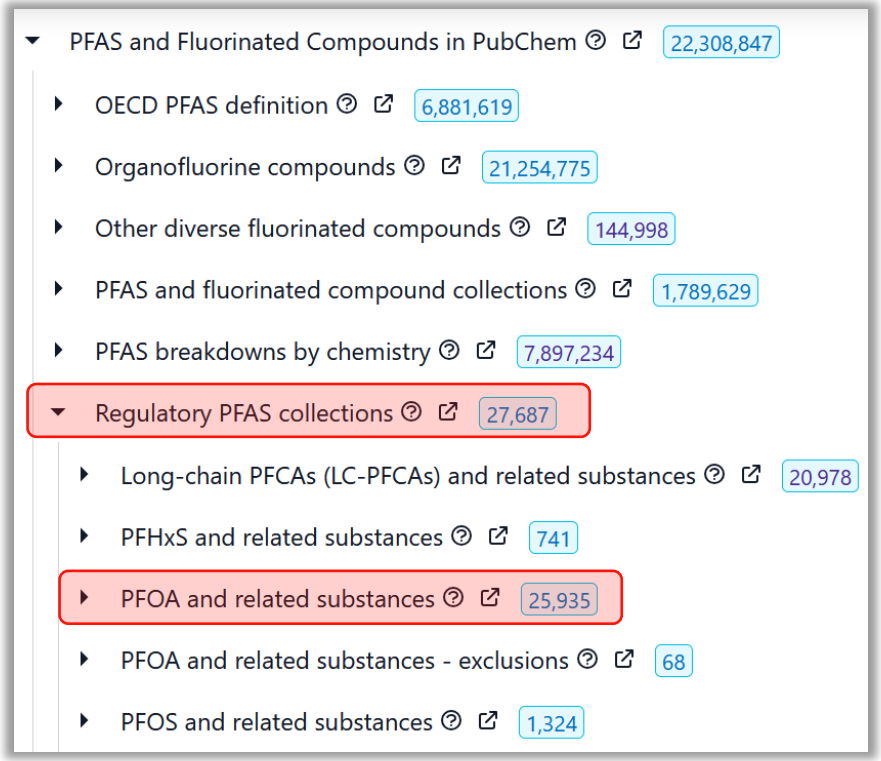
Schymanski *et al*. (2021) DOI: 10.1186/s13321-021-00489-0

Exploring Chemicals in PubChem



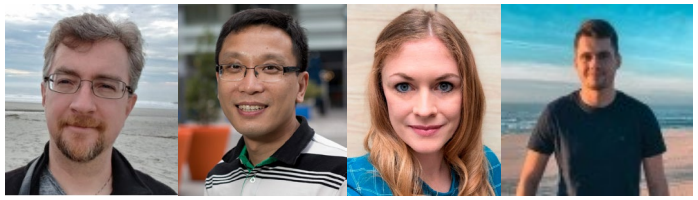


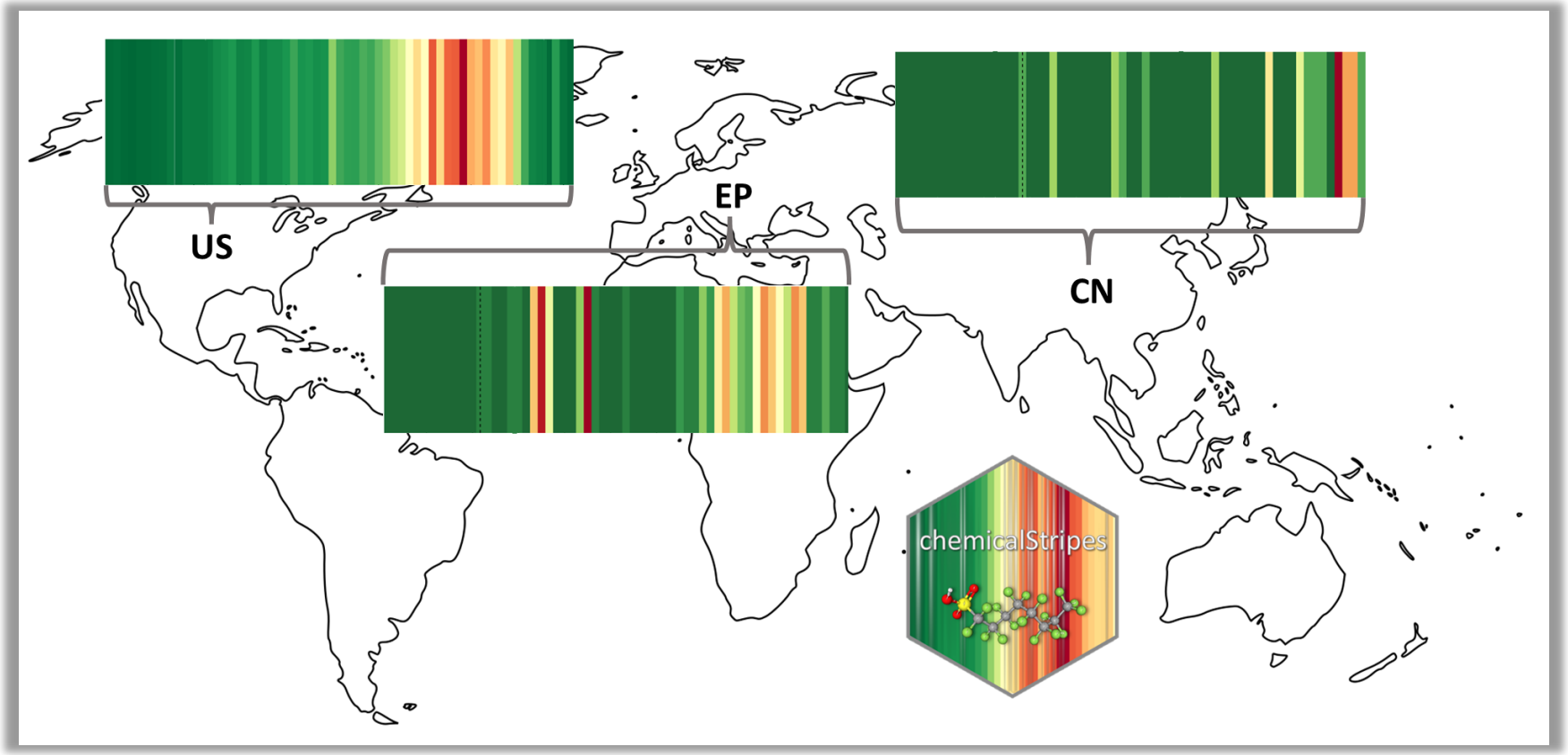




Schymanski *et al.* (2021) DOI: 10.1186/s13321-021-00489-0

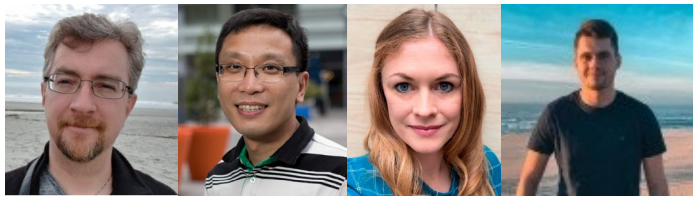
Regional Analysis of Patent Information

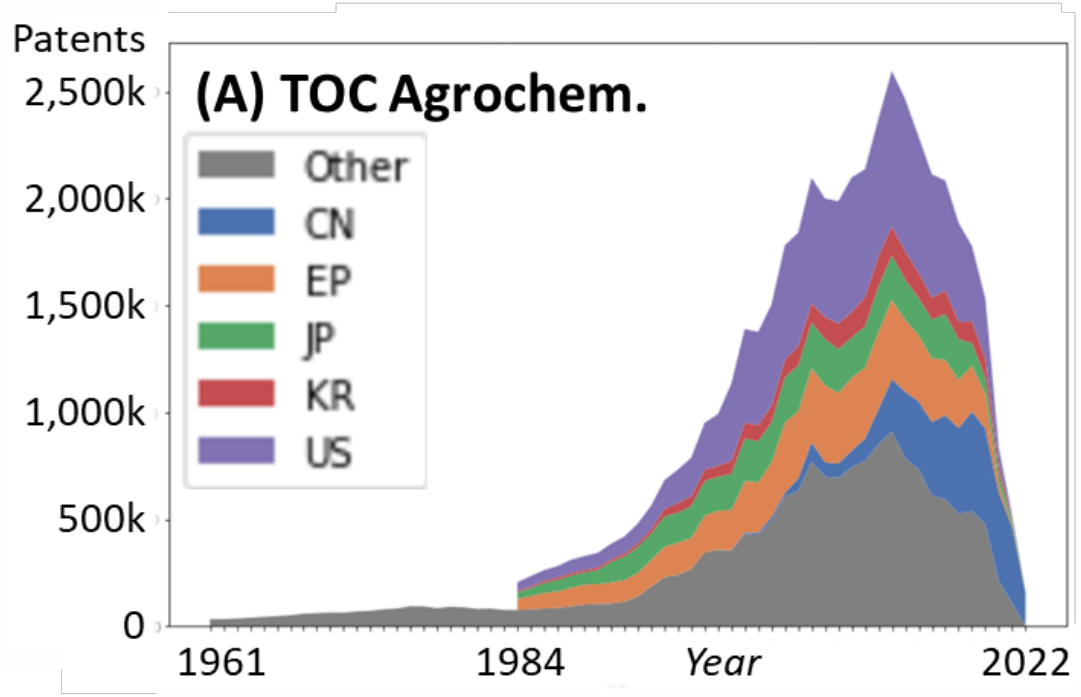


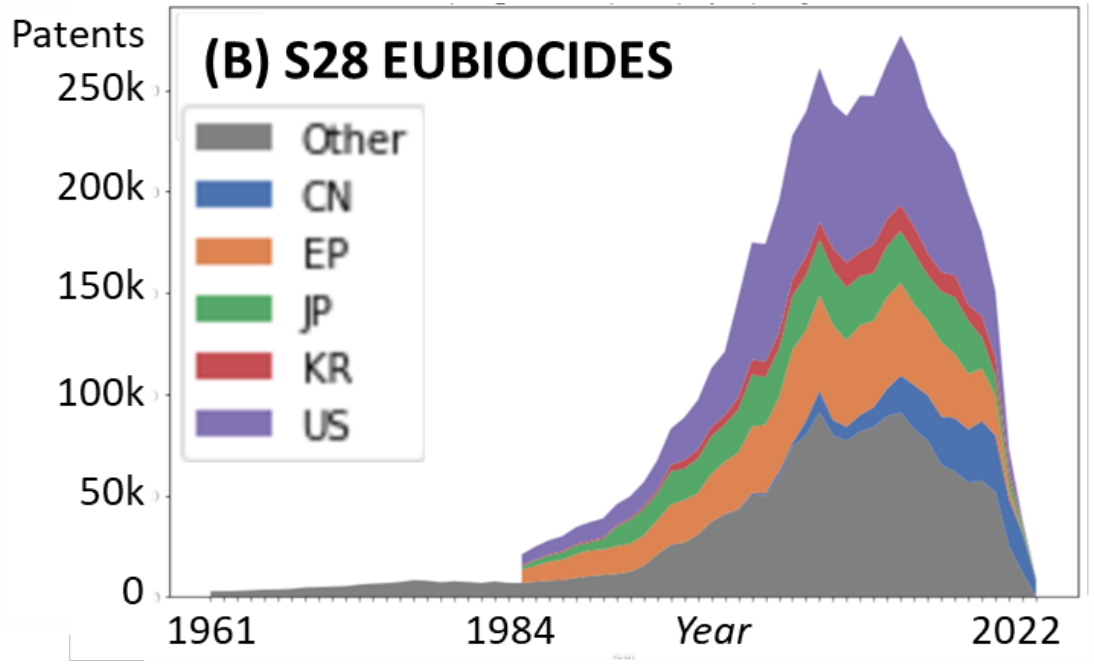


Aurich *et al.* (2024) ES&T Letters DOI: 10.1021/acs.estlett.4c00560
https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

Trends in Agrochemicals and Regions

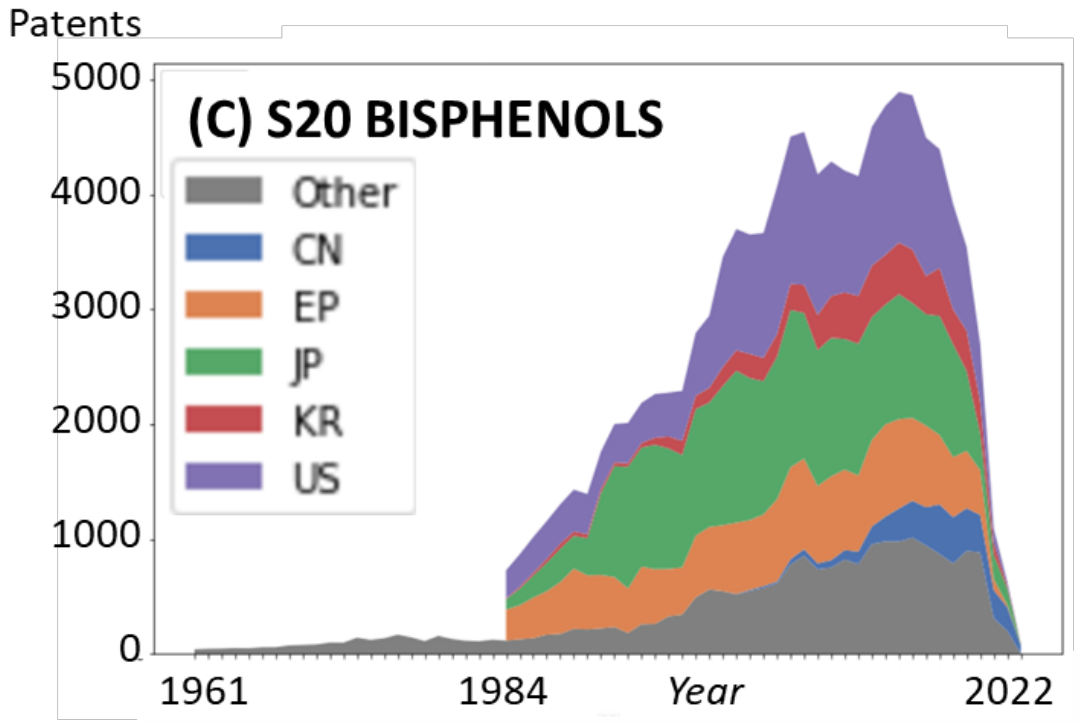


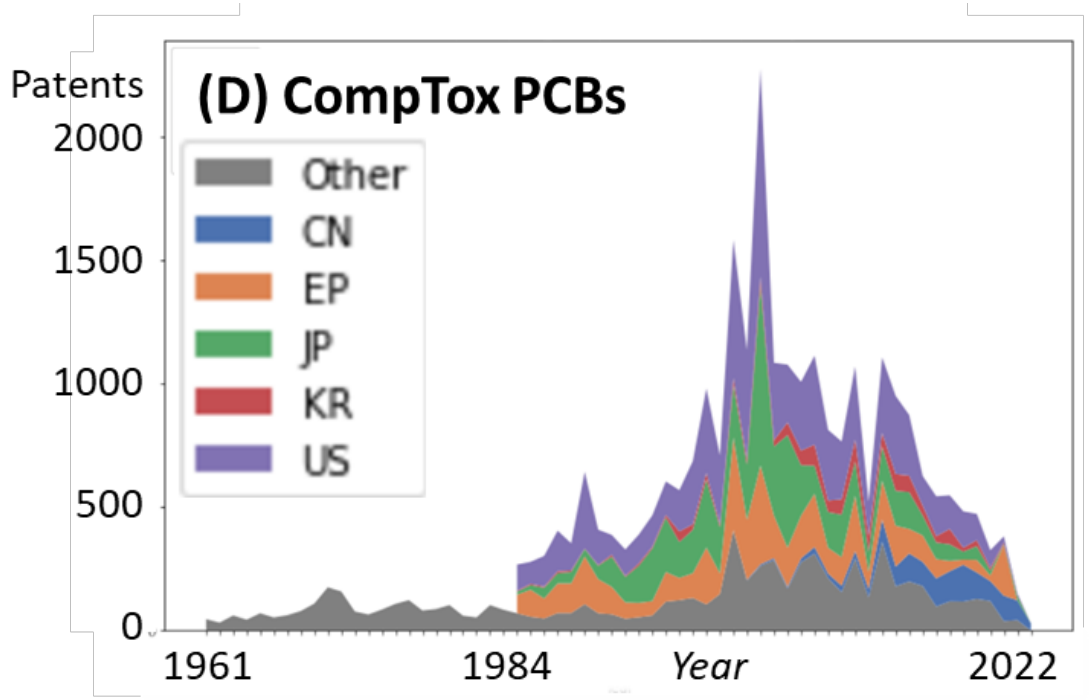




Aurich *et al.* (2024) ES&T Letters DOI: 10.1021/acs.estlett.4c00560
https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

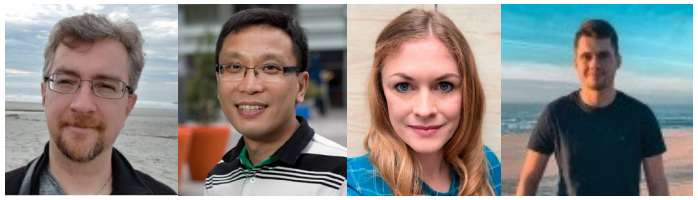
Trends in Bisphenols and Polychlorinated Biphenyls (PCBs)

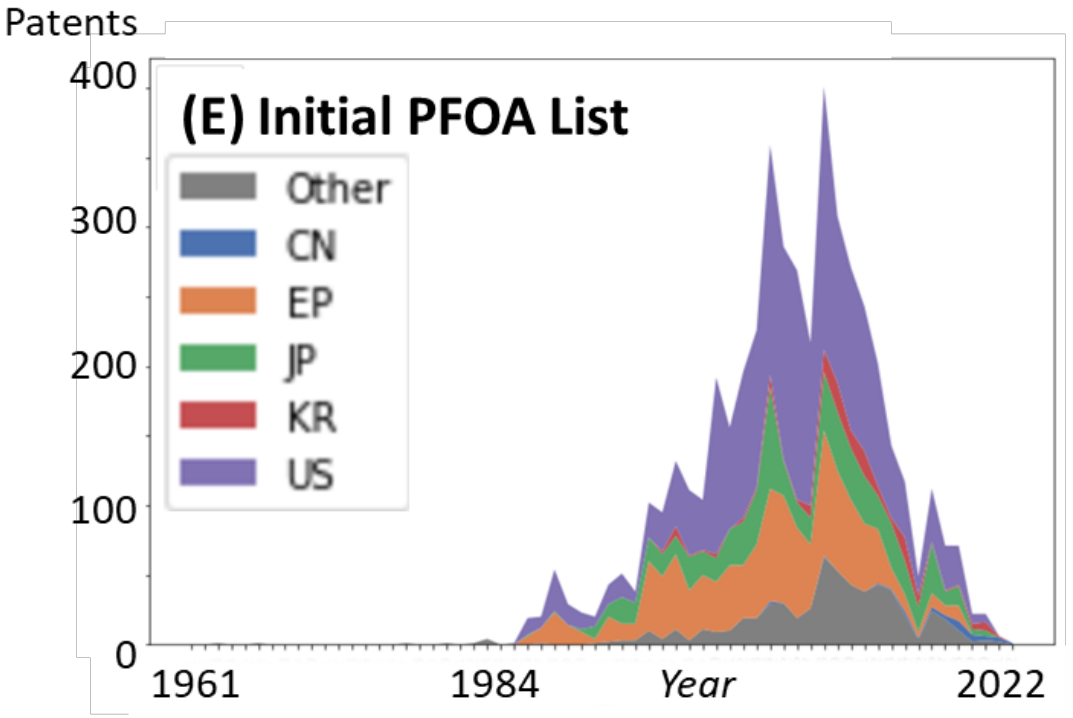


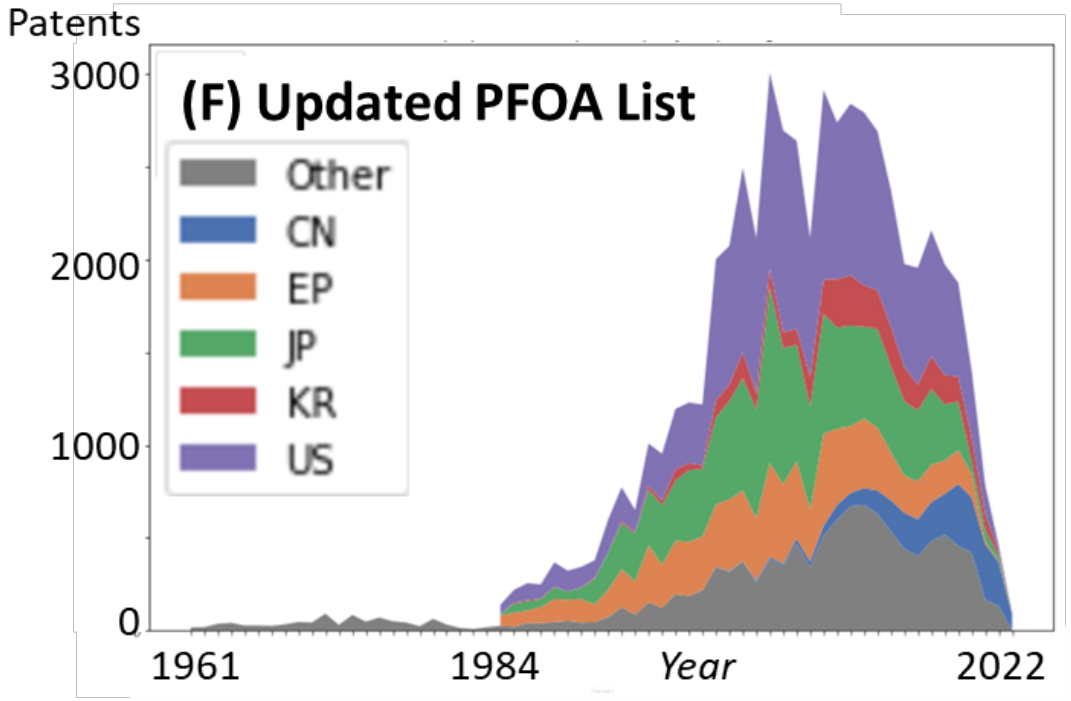


Aurich *et al.* (2024) ES&T Letters DOI: 10.1021/acs.estlett.4c00560
https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

Trends in PFAS

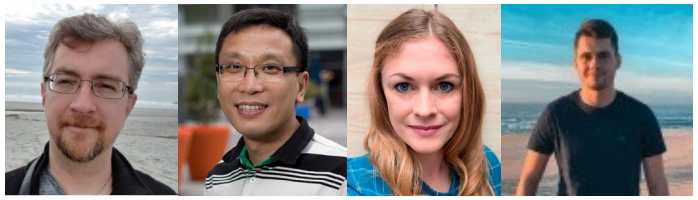


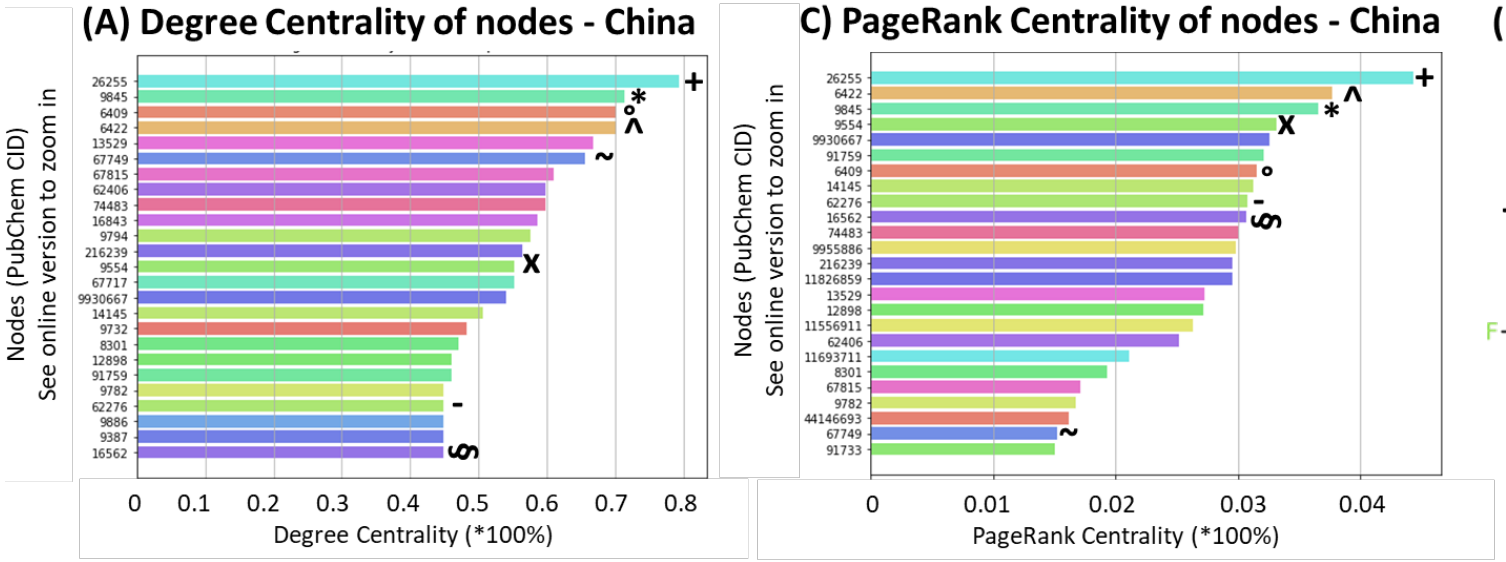


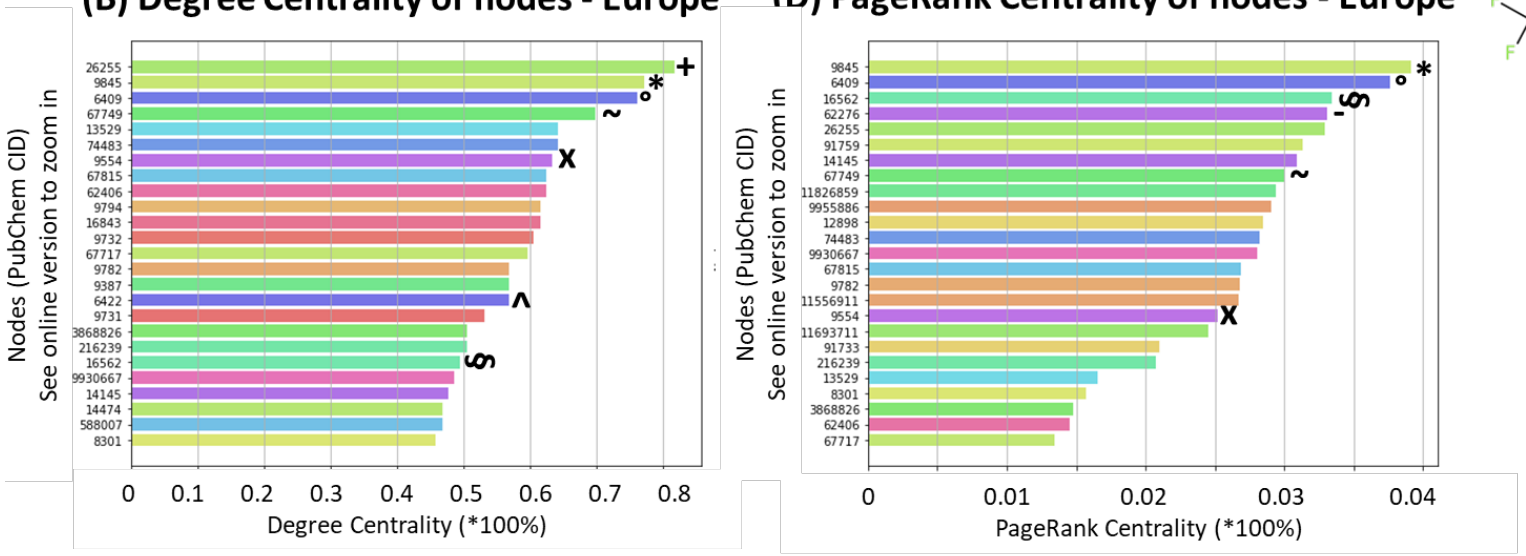


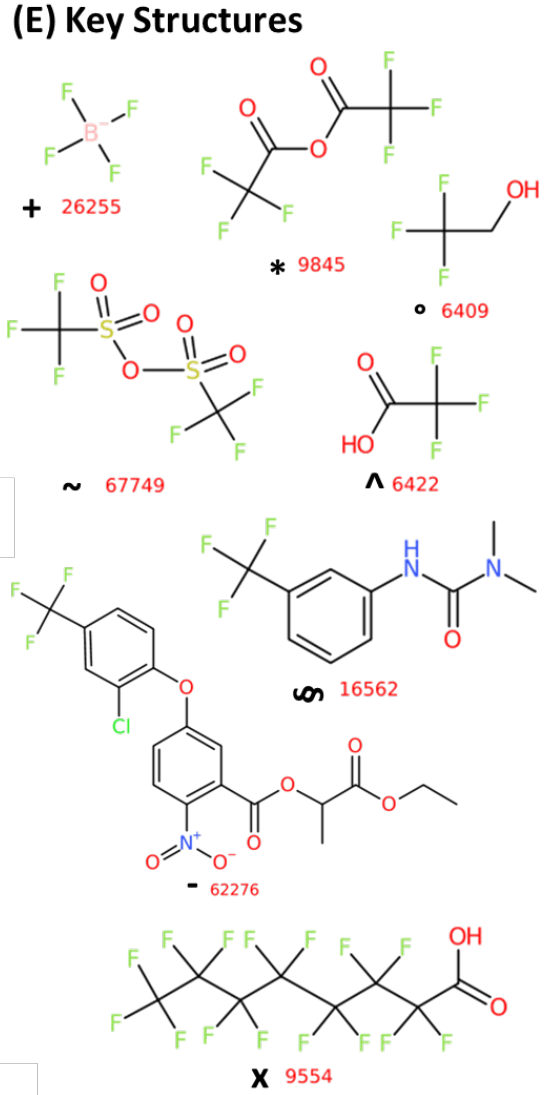
Aurich *et al.* (2024) ES&T Letters DOI: 10.1021/acs.estlett.4c00560
https://gitlab.com/uniluxembourg/lcsb/eci/chemicalstripes

Which are the “up and coming” PMT PFAS?



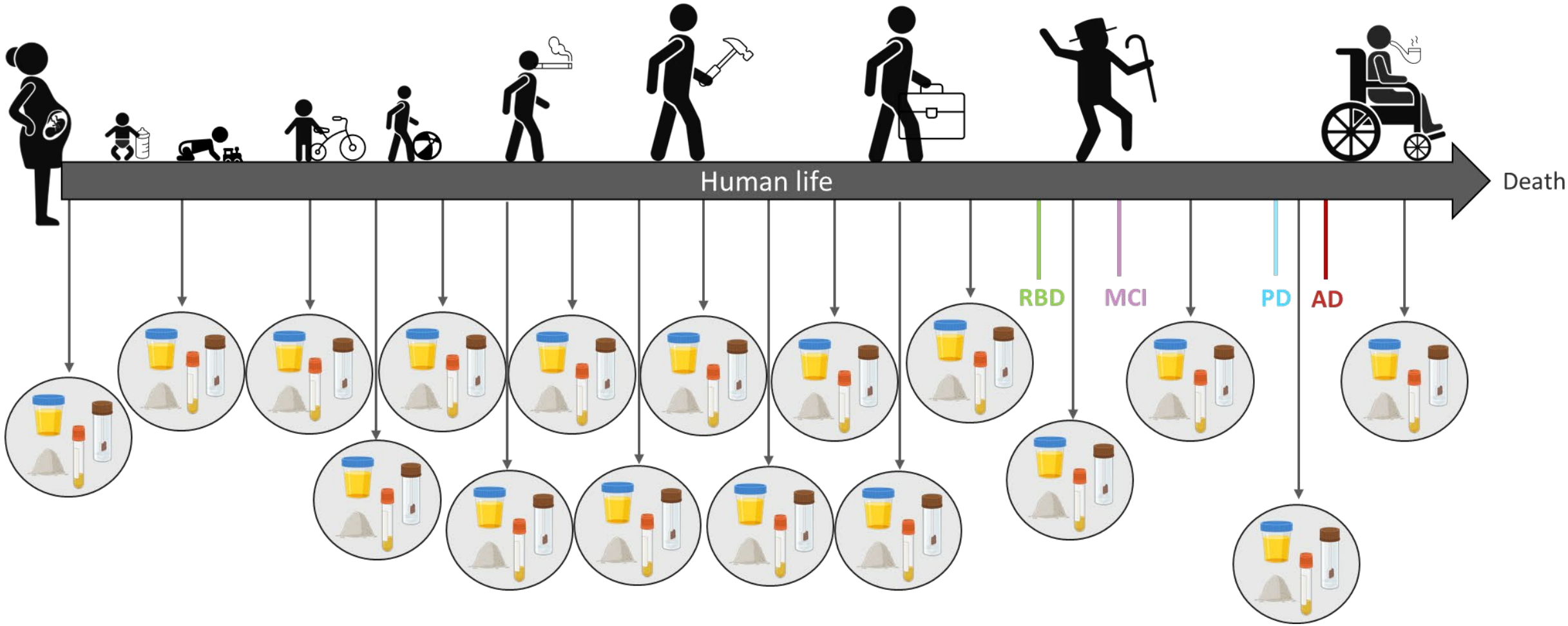






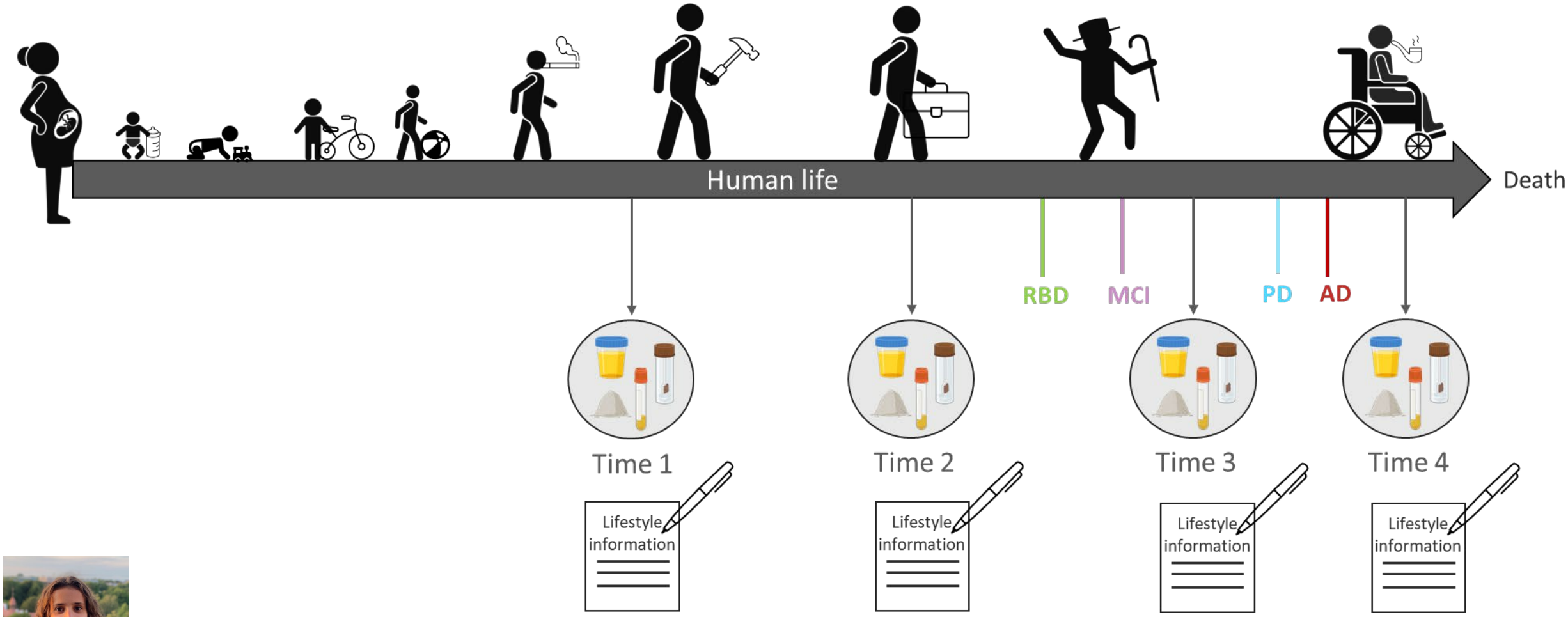
Aurich *et al.* (2024) ES&T Letters DOI: [10.1021/acs.estlett.4c00560](https://doi.org/10.1021/acs.estlett.4c00560)

The Exposomics Dream - Wouldn’t it be nice if …



Source: Begoña Talavera Andújar, PhD Thesis.

The Exposome Reality – Making the best of the data we have



Source: Begoña Talavera Andújar, PhD Thesis.

Merci/Danke schön/Thank you!

Email: emma.schymanski@uni.lu
@ESchymanski / @schymane@mstdn.social
https://wwwen.uni.lu/lcsb/research/environmental_cheminformatics/
https://gitlab.com/uniluxembourg/lcsb/eci/

Today's slides

THE GOVERNMENT OF THE GRAND DUCHY OF LUXEMBOURG
Ministry of the Environment, Climate and Sustainable Development

Funded by the European Union
H2020: 101036756

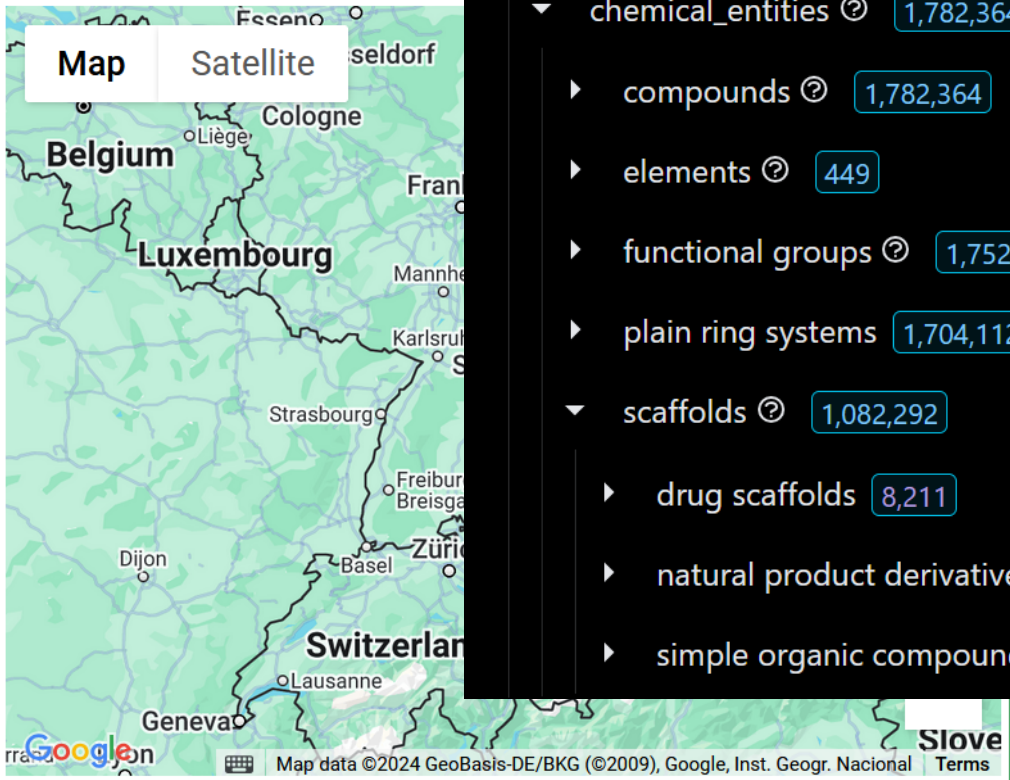
Luxembourg National Research Fund

PS: New Developments In Action

MolGenie - Molecular Design and Engineering

MolGenie is developing and applying novel compound 2D and 3D design technologies aiming to support new
follows function" - this Louis Sullivan principle guides compound design and engineering. Using these rules an
proprietary structure-property-relationship datasets they create intelligent molecules that realize desired funct

Organization	MolGenie GmbH
Category	Research and Development
URL	http://molgenie.com/
License Note	CC-BY 4.0
License URL	https://creativecommons.org/licenses/by/4.0/
Contact Name	Lutz Weber
Address	Felix-Dahn-Str. 4, Stuttgart, Germany, 70597
Data Source ID	27441
Data in PubChem	1,797,833 Live Substances
Last Updated	2024/09/10



Browse MolGenie Ontology

- MolGenie Ontology 1,782,364
  - chemical_entities 1,782,364
    - compounds 1,782,364
    - elements 449
    - functional groups 1,752,097
    - plain ring systems 1,704,112
    - scaffolds 1,082,292
      - drug scaffolds 8,211
      - natural product derivatives 184,221
      - simple organic compound scaffolds 993,067

https://pubchem.ncbi.nlm.nih.gov/source/27441