



Blurred lines: Crossing the boundaries between the chemical exposome and the metabolome

Cristina Balcells¹, Yitao Xu¹, Rubén Gil-Solsona²,
 Léa Maitre^{3,4,5}, Pablo Gago-Ferrero² and Hector C. Keun¹

Abstract

The aetiology of every human disease lies in a combination of genetic and environmental factors, each contributing in varying proportions. While genomics investigates the former, a comparable holistic paradigm was proposed for environmental exposures in 2005, marking the onset of exposome research. Since then, the exposome definition has broadened to include a wide array of physical, chemical, and psychosocial factors that interact with the human body and potentially alter the epigenome, the transcriptome, the proteome, and the metabolome. The chemical exposome, deeply intertwined with the metabolome, includes all small molecules originating from diet as well as pharmaceuticals, personal care and consumer products, or pollutants in air and water. The set of techniques to interrogate these exposures, primarily mass spectrometry and nuclear magnetic resonance spectroscopy, are also extensively used in metabolomics. Recent advances in untargeted metabolomics using high resolution mass spectrometry have paved the way for the development of methods able to provide in depth characterisation of both the internal chemical exposome and the endogenous metabolome simultaneously. Herein we review the available tools, databases, and workflows currently available for such work, and discuss how these can bridge the gap between the study of the metabolome and the exposome.

Addresses

¹ Institute of Developmental and Reproductive Biology (IRDB), Division of Cancer, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, UK

² Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Barcelona, Spain

³ ISGlobal, Barcelona, Spain

⁴ Universitat Pompeu Fabra (UPF), Barcelona, Spain

⁵ CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

Corresponding authors: Balcells, Cristina (c.balcells-nadal@imperial.ac.uk); Keun, Hector C. (h.keun@imperial.ac.uk)

Current Opinion in Chemical Biology 2024, 78:102407

This review comes from a themed issue on **Omics - Metabolomics (2023)**

Edited by **James McCullagh** and **Hector Keun**

For complete overview of the section, please refer the article collection - **Omics - Metabolomics (2023)**

Available online 11 December 2023

<https://doi.org/10.1016/j.cbpa.2023.102407>

1367-5931/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords

Metabolomics, HRMS, Exposomics, Chemical exposome, Mass spectrometry.

Introduction

The *exposome* is defined as the aggregation of all contacts between external exposures and the human body throughout its lifetime [1–3]. This definition encompasses exposure to diet and dietary supplements, medicinal and recreational drugs, personal care and consumer products (PCCPs), air and water pollution [4], physical activity, noise, radiation, as well as non-physical factors such as socioeconomic status [5], access to green [6] or blue spaces [7], or perception of one's surroundings [8]. Since the exposome concept was coined by Wild in 2005 [1], we have witnessed *exposomics* maturing as a scientific field and becoming a new ‘-omics’ layer in its own right [9–11]. The paradigm on which exposomics is erected relies fundamentally on three main pillars (Figure 1a): (1) Measuring the exposome (with a variety of tools ranging from mass spectrometry-based approaches, geospatial models, portable sensors, to questionnaires [10]); (2) Measuring human adverse health outcomes potentially caused by environmental exposures (at any scale from precision medicine to public health); and (3) Statistically evaluating the associations between exposure signatures and adverse health outcomes, and identifying the molecular mechanisms or pathways driving exposure-disease associations. The latter can be achieved by strategies that include the definition of adverse–outcome pathways (AOPs) [12] and the integration with other -omics layers [13,14].

Exposome research has revealed environmental exposures as key contributors to major human diseases and health conditions, including cancer [15,16], immune-mediated diseases [17,18], cardiometabolic health [19], hepatic health [20] or mental health [7]. In addition to studies considering disease prevention in the general population [19], sub-populations with abnormal

Abbreviations

AOP	Adverse outcome pathway
EDC	Endocrine-disrupting chemical
ExWAS	Exposome-wide association study
GC	Gas chromatography
HRMS	High-resolution mass spectrometry
LC	liquid chromatography
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MWAS	Metabolome-wide association study
NMR	Nuclear magnetic resonance
PCA	Principal component analysis
PCCP	Personal care and consumer products
PCDD	Polychlorinated dibenzodioxins
PFAS	Perfluoro-alkylated substances
QqQ	Triple-quadrupole
VOC	Volatile organic compounds

exposure or disease risk profiles have been studied in exposome research e.g., those with geographical [21,22] or occupational [23,24] exposure or during pregnancy and early life [25,26], due to the crucial and longer-term impact that chemical exposures can have during these vulnerable periods. In this regard, extensive data on chemical exposure, omics, and health outcomes are being collected in multiple large-scale mother-child longitudinal cohorts and research infrastructures across the globe [25–27], including the characterization of pregnancy-related biofluids such as cord blood [28], amniotic fluid [29], or breast milk [30–33].

A significant subset of the exposome is what we often refer to as the *chemical exposome*. This subset encompasses all chemical species (and their associated transformation products), either from biological or synthetic origin, capable of entering the human body through different pathways such as ingestion, inhalation, or dermal absorption. These molecules can arise from sources such as diet, drugs, and dietary supplements, PCCPs, as well as water and airborne substances. The most suitable techniques to measure these chemicals in human tissues or biofluids involve the use of LC (liquid chromatography) and GC (gas chromatography), coupled to either tandem or high-resolution mass spectrometry (MS/MS or HRMS, respectively), and nuclear magnetic resonance (NMR) spectroscopy. Importantly, these techniques constitute the primary methodologies used in metabolomics, defined as the holistic measurement of any small molecules (<2 kDa) present in a living system, regardless of origin [34]. As reviewed elsewhere [3,34], a certain chemical space and semantic overlap exists between both omics, with the metabolomics and exposome research communities referring to a common set of molecules with different semantics and research focus angle (Figure 1b). For this

reason, and particularly with the emergence of untargeted approaches, the analytical boundaries between metabolomics and chemical exposomics are rapidly blurring.

Can we simultaneously measure the endogenous metabolome and the internal chemical exposome effectively?

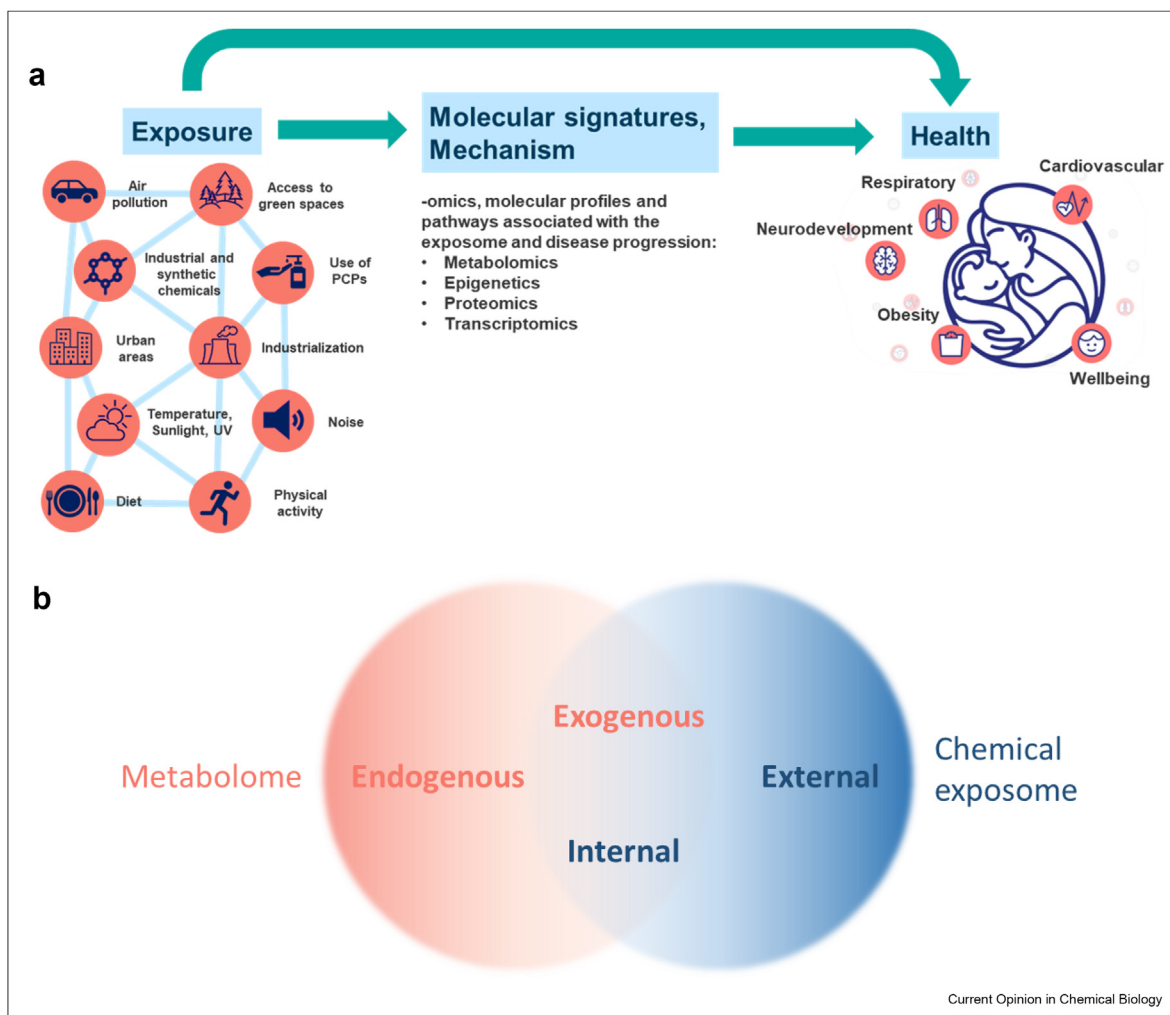
Untargeted metabolomics workflows designed to identify the metabolic signatures of diseases such as cancer [35], cardiometabolic health and obesity [36], or neurodegeneration [37] often deliberately filter out any detected features that are not of a clear human endogenous origin. Even if retained, these often remain unannotated because the libraries used for compound identification in metabolomics are biased towards the endogenous human metabolome whereas the food metabolome and xenobiotics are often under-represented. Indeed, the endogenous metabolome of the most analysed human biospecimens such as blood [38] or urine [39] has been described in detail, making the exclusion of unknowns or exogenous compounds straight-forward.

Whereas untargeted methods are more prevalent in the metabolomics research field and targeted metabolomics is mostly prevalent in clinical settings, exposomics research remains split more evenly between highly targeted methods (Table 1), broad scope targeted with multiple analyte families per assay (Table 2) and fully untargeted approaches (Table 3). Targeted, quantitative methods are still widely used when measuring the chemical exposome, due to need for reliable biomonitoring data and the desire to define quantitative exposure-adverse health outcome associations for regulatory authorities and policymakers. Numerous governmental agencies or national laboratories worldwide publish exposure biomonitoring data of their populations periodically [40,41].

Most targeted analytical methods employed to characterize the exposome are tailored to distinct groups of established exposures (as outlined in Table 1). Typically, these methods measure either single or up to a few 10's of compounds. Such methods are usually validated for a specific biofluid or matrix. Notably, most of the MS instrumentation and chromatographic columns shown in Table 1 are routinely used in metabolomics studies [42], and as MS instrumentation improves, broader targeted methods have been reported that can simultaneously measure different families of exposures in different matrices ranging from >50 analytes [31] to hundreds of analytes [43] (Table 2).

This expansion of coverage now includes suspect screening [71,72], based on specific families of exposures (e.g., pesticides [72], endocrine-disrupting

Figure 1



The relationship between the exposome, the metabolome and human health.

chemicals (EDCs) [29] or PFAS [24,28]), and completely untargeted approaches [5]. These methods are becoming crucial to identify and characterize previously unknown exposures, exposure families and exposure metabolites.

Notably, a major proportion of exposome studies report endogenous metabolism and exogenous chemicals in human biofluids in separate assays [13,73–76], either targeted [13] or untargeted (Table 3). In some cases, either the endogenous or the exogenous [5,77] information is deliberately excluded from the reported results, with the potential concomitant loss of mechanistic insight into the biological response to exposures. The main motivations for opting for separate assays may be the existence of previously established chromatographic methods in the metabolomics or environmental

analytical chemistry fields, fields with traditionally little overlap or interaction, or the inclusion of GC-based approaches to increase the coverage for exogenous chemicals [78] (Table 1). However, an increasing number of studies are showing promising results that pave the way for the joint endogenous-exogenous venture in a single assay [14,27,75,77,79], as reviewed in the next sections.

The combination of metabolome and chemical exposome determination in a single analytical run or workflow is gaining ground in untargeted studies [14,22,27,77,79] (Table 3), but also in targeted studies [31,43] (Table 2). Table 3 also shows a clear trend in the community of combining HRMS assays with complementary chromatographic approaches such as GC and LC, or LC reversed-phase and HILIC. We believe

Table 1

Examples of targeted assays for established families of chemical exposures.

Main source	Exposure family	Matrix	Chromatography	MS	LOD
PLASTICIZERS	Phthalates [44]	U	LC Chromolith Flash RP-18e	Thermo TSQ Quantum QqQ MS	0.2–0.4 ng/ml
NATURAL PRODUCTS	Mycotoxins [30]	BM	LC Waters ACQUITY HSS T3	Sciex 6500	0.005–4.6 ng/ml
	Phytoestrogens [45]	S	LC Synergi Polar-RP 2.5 micron	Sciex API 4000	2 ng/ml (LOQ)
COMBUSTION PRODUCTS	Polycyclic Aromatic Hydrocarbon Metabolites [46]	U	LC Chromolith High Resolution RP-18	Sciex 5500 and 6500 QqQ MS	0.008–0.09 ng/ml
TOBACCO SMOKE	Tobacco-specific nitrosamines (TSNA) [47]	U	LC Phenomenex Gemini-NX C18	Sciex API 5500 QTRAP	0.4–10 pg/ml
	Cotinine [48]	S, SV	LC Waters Symmetry C18	Sciex API III + QqQ MS	NA
	Heterocyclic Aromatic Amines [49]	U	LC Agilent Zorbax Eclipse Plus C18	Sciex API 5500 QTRAP system	0.31–0.83 pg/ml
PESTICIDES, INSECTICIDES, AND INSECT REPELLENTS	Pyrethroids [50]	U	LC Thermo Betasil C18	Thermo Quantum Ultra QqQ MS	0.03–0.4 ng/ml
	Organochlorine Pesticides and Metabolites [32]	BM	GC DB-35MS	Agilent 7000 MS	0.066–1.8 ng/kg wet weight
	Dialkyl phosphate pesticides [51]	U	LC Waters ACQUITY BEH C18	Waters Xevo GS-2 QTOF MS	0.15–1.2 ng/ml
	Other organophosphate insecticides [52]	U	LC Thermo Betasil phenyl	Sciex API 4000	0.1–1.5 ng/ml
	Carbamate Insecticides [53]	B, U	LC Waters ACQUITY BEH C18	Waters Xevo TQD MS	0.21–0.56 ng/ml in urine 0.41–0.44 ng/ml in blood
	Insect Repellent and Metabolites (DEET) [54]	U	LC Phenomenex Prodigy 5u Phenyl-3 (PH-3)	Thermo TSQ Quantum Ultra™ QqQ MS	0.1–1 ng/ml
HERBICIDES	Atrazine and metabolites [55]	U	2D-HPLC Phenomenex Luna SCX, Gemini C6-Phenyl	Thermo TSQ Quantum Ultra MS/MS	10–50 µg/l
	2,4-D, 2,4,5-T and metabolites [52]	U	LC Thermo Betasil phenyl	Sciex API 4000 QTRAP	0.1–0.2 ng/ml
FUNGICIDES	Sulfonylurea and metabolites [56]	U	LC Phenomenex Prodigy column	Thermo TSQ 7000 QqQ MS	0.125–1 ng/ml
	Fungicides and metabolites [53]	B, U	LC Waters ACQUITY BEH C18	Waters Xevo TQD MS	0.17–11.8 ng/ml in urine 0.32–12.3 ng/ml in blood
FLAME RETARDANTS	Urinary Flame Retardants [57]	U	LC Agilent ZORBAX Eclipse XDB-C8	Sciex 5500 QTRAP MS	0.05–0.16 ng/ml
	PBDEs and PBB 153 [58]	S	GC DB5HT	Thermo MAT 95XP	0.1–1.2 ng/g lipid (LOQ)
Industrial and PCCP	PFAS [59]	S, U	LC Chromolith HR RP-18e	Sciex QqQ 5500 and QTRAP 6500 MS	0.1 ng/ml
	Parabens [60,61]	U	LC Betasil C18 column	Waters QqQ	0.02–0.06 ng/ml
		F	LC ACQUITY BEH C18	Waters Quattro Premier XE QqQ	10 µg/l (LOQ)
MULTIPLE SOURCES	Bisphenols [62]	U	LC Chromolith HR RP-18e	Sciex 5500 QTRAP MS	0.03–1 ng/ml
	VOCs and metabolites [63,64]	B/S/P	GC DB-624	Thermo MAT 95XP magnetic sector MS	0.01 µg/l
		U	LC Waters ACQUITY HSS T3	Sciex API 5500 Triple Quad	2.5–9.85 ng/ml
	PCBs [65,66]	–	Ion mobility	6560 IMS-QTOF MS	NA
		S	GC DB-5MS	Thermo MAT95XL	NA
	PCDDs and Dibenzofurans [67]	S	GC DB-5MS capillary column	Thermo DFS HRMS	0.005–9 ng/kg wet weight

Abbreviations: PBDE (Polybrominated Diphenyl Ethers), PCB (Polychlorinated Biphenyls), PCCP (Personal Care and Consumer Products), PCDD (Polychlorinated dibenzodioxins), PFAS (Perfluoro alkylated substances), VOC (Volatile organic compounds). Matrices – B: Blood, BM: Breast milk, F: Food, P: Plasma, S: Serum, SV: Saliva, U: Urine. Thermo can refer to ThermoFisher or its associated company ThermoFinnigan. This table depicts a selection of exposure classes that are included in the most relevant environmental biomonitoring initiatives and is by no means comprehensive.

Table 2

Examples of targeted multi-class studies reporting different families of chemical exposures and the metabolome in a single LC-MS assay. This table depicts a selection of a few representative examples of targeted studies that capture the metabolome and/or the chemical exposome in a single analytical assay and is by no means comprehensive [43,52,67–70].

	Endogenous Metabolome	Food	Pharmaceuticals	Phthalates	Mycotoxins	Phytoestrogens	PAH and metabolites	TSNA	Cotinine	Heterocyclic aromatic amines	Pyrethroids	Organochlorine pesticides	Dialkyl Phosphates	Other organophosphates	Carbamate Insecticides	DEET and metabolites	Atrazine and metabolites	2,4-D, 2,4,5-T and metabolites	Sulfonylurea and metabolites	Fungicides and metabolites	Urinary flame retardants	PBDEs and PBB153	PFAS	Parabens	Bisphenols	VOC and metabolites	PCBs	PCDDs and Dibenzofurans
Zhang, J Sep Sci. 2017 [67]																												
Hung et al. Appl. Sci. 2019 [68]																												
Olsson. Anal. Chem. 2004 [52]																												
Aly et al. Anal. Bioanal. Chem 2022 [69]																												
Flasch, JACS 2022 [70]																												
González-Domínguez, Anal. Chem 2020 [43]																												

this will continue to be the case in the future, particularly as metabolomics and chemical exposomics converge, and can be further expanded with multiple reversed-phase columns with different chemistries or even different MS fragmentation modes (e.g., collision-induced dissociation vs. electron-activated dissociation). The use of multiple analytical assays is a decision that should be carefully considered in the planning phase of a project, depending on both scientific factors (research focus and aims, technical expertise, need to harmonize the produced data with existing datasets) and logistical factors (time, instrument, sample and budget availability).

Figure 2 shows the comparison of the chemical space occupied by a widely used definition of the metabolome (using the human metabolome database, HMDB) with that of different families of environmental exposures. This figure illustrates that most families of exposures possess similar chemical properties to the metabolome and should therefore become increasingly detectable as HRMS instruments become more sensitive. The sensitivity required is unique to each individual chemical exposure and will be sufficient once an assay manages to capture the range of concentrations in which an individual exposure has no effect on human health outcomes. In parallel, other advances such as ion mobility [65], composite stationary phases [81] or advanced sample preparation strategies [82] can also contribute to expanding the chemical space explored by metabolome/chemical exposome studies.

Simultaneous measurement of the metabolome and the chemical exposome

Simultaneous measurement of the metabolome and the chemical exposome: data acquisition

In the last 5 years, studies that combine both omics in one analytical run are becoming more prevalent in the literature [14,27,77,79]. Chen et al. [77] use a common reversed-phase LC-HRMS assay to analyse the organic exposome and the metabolome in urine of >1000 pregnant mothers in the first and third trimester of pregnancy, establishing longitudinal associations between the exposome in the first trimester and the metabolome in the third. Panagopoulos et al. [27] reported an untargeted LC-QTOF MS analysis of maternal and cord blood in the CIOB pregnancy cohort, annotating both endogenous and exogenous features and reporting cross-sectional associations between both sets. They also report a useful strategy to automate the classification of endogenous and exogenous features using the HUMANBLOOD database. Gao et al. [14] integrate multiple omics and exposure profiles for a single individual, reporting the endogenous metabolome and chemical exposures in plasma from a HILIC + RP-HRMS dataset. All three studies opt for MS¹-based approaches, obtaining MS² data only from pooled quality control samples in the case of Gao et al. [14] or for selected compounds in the case of Panagopoulos et al. [27]. This limits the confidence with which the identity and presence of exogenous compounds can be reported to only those present in a significant number or the majority of the samples in a given study. Addressing this

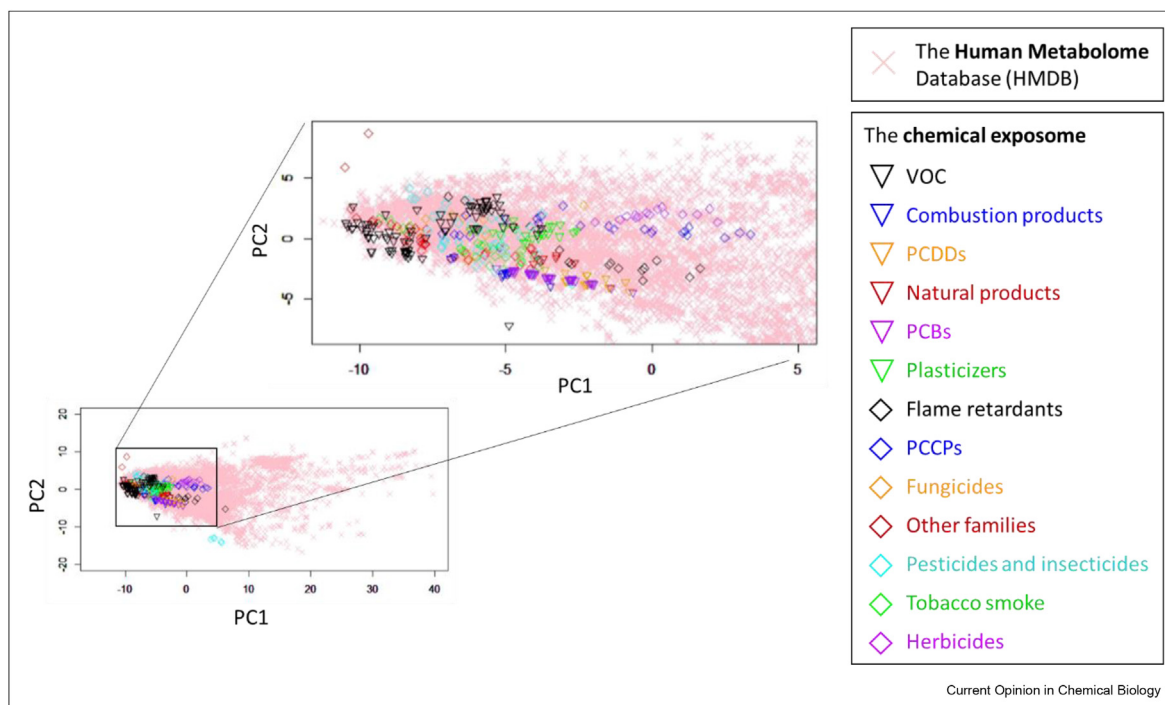
Table 3

Representative untargeted LC-HRMS studies including metabolome and chemical exposome measurements.

Study description	Exposome chemicals	N (MS level 1)*	Matrix	Exogenous	Endogenous
Panel study of the effect of metals on commuters [22]	In-vehicle, particulate matter exposures (Metal: Al, Pb, and Fe; PM2.5 mass, black carbon, particle-bound polycyclic aromatic hydrocarbons (pb-PAHs))	–	P	LC-Q-Exactive	
Panel study of the effect of exposure to swimming pool disinfection by-products [75]	Disinfection by-products	13	S	LC-Agilent 6550 QTOF	
Metabolic profiling of participants in the Michigan PBB Registry [74]	PCB-153 and PBB-153	120	P	GC–MS	– LC-Q-Exactive
Early-life study with the effect of PFAS exposure in the metabolome of NAFLD patients [73]	PFAS	–	P	(RP ESI–) LC-Q-Exactive HF	(HILIC ESI+) LC-Q-Exactive HF
Pregnancy cohort [27]	Parabens, PFAS, phthalates, cyclic volatile methylsiloxanes, food-derived methylxanthines	19	CB, B	LC-Agilent 6550 QTOF	
Case/control study with cholangitis patients (PSC/ PBC) [76]	Parabens, PFAS, carbamate insecticides, organophosphate insecticides, herbicides, fungicides	126	P	GC-Q-Exactive	(HILIC + RP) LC-FTMS
Longitudinal multi-omics integration with the personal exposome of a single individual [14]	Bisphenols, flame retardants, dioxins, carcinogens, mycotoxins, PFAS, pesticides, pharmaceuticals, PCCPs, plasticizers, PCBs, polychlorinated naphtalenes, PAH	15	P	HILIC–Q-Exactive Plus LC-Exactive	RP-Q Exactive –
Panel study of smokers/non-smokers metabolome and chemical exposome association [79]	Nicotine, cotinine, 2-HMPA, 3-HPMA, HMPMA, CEMA, DHBMA, phenylglyoxylic acid, 3-benzylpropionic acid	9**	U	LC-tripleTOF	
Pregnancy cohort [77]	Pesticides (organochlorine, organophosphate, carbamate, pyrethroid, triazine and amide), antimicrobial agents, phthalates, phenols, parabens, nicotine and related, perfluorochemicals, plasticizers and fire retardants and plant metabolites	198	U	LC-Q-Exactive	
Development of a dietary exposome library and evaluation of the role of poly(phenols) in reference human samples [80]	Aromatic ketones, benzoic acids, ellagic acids, caffeoylquinic acids, catecholamines, coumarins, hippuric acid, hydroxytoluenes, phenylamines, stilbenes, urolithins, valerolactones, and xanthonoids	66/123 ***	S, U	LC-Q-Exactive HF-X	–

Matrices – B: Blood, BM: Breast milk, CB: Cord blood, F: Food, P: Plasma, S: Serum, SV: Saliva, U: Urine. Abbreviations: CEMA: N-Acetyl-S-(2-carboxyethyl)-L-cysteine; DHBMA: N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine; HMPA: Hydroxypropyl-mercaptopuric acid; HMPMA: N-Acetyl-S-(3-hydroxy-1-methylpropyl)-L-cysteine; RP: Reversed-phase; PBC: primary biliary cholangitis; PSC: primary sclerosing cholangitis. Thermo can refer to ThermoFisher or its associated company ThermoFinnigan. * Within this table, we have considered metabolomics standards initiative (MSI) Level 1 annotations all compounds that have been validated with injected reference standards in terms of at least *m/z* and retention time. In some **publications**, there is no information of whether this process involved confirmation of the MS/MS spectrum of the compounds. ** This study only aims to validate the identification to MSI Level 1 for features with significant differences between the study groups. *** 66 detected on biofluids, 123 on the spectral library generated by this study. This table depicts a selection of a few representative examples of studies featuring untargeted HRMS for the metabolome and/or the chemical exposome and is by no means comprehensive. Due to the inherent chemical heterogeneity of untargeted HRMS datasets, the chemical exposures column is not comprehensive **and** provides highlights of the chemical exposure classes identified in each study. In some of the studies, some chemical exposures classes were measured in parallel using targeted methods.

Figure 2



Comparing the chemical space of the human metabolome to the chemical exposome. Principal component analysis (PCA) using the chemical descriptors (CD) calculated using the Chemistry Development Kit (CDK) framework on all entries of the human metabolome database (HMDB) [83] and a list of 350 exposome chemicals part of the CDC biomonitoring list [40]. Calculation of the CDs and PCA plot were performed using functions provided with the Retip R package [84].

issue, Frigerio *et al.* [79] recently reported a novel data-dependent acquisition tripleTOF-MS/MS method for the simultaneous measurement of endogenous and exogenous compounds in urine samples, combining reversed-phase and HILIC. Approaches with MS/MS on every sample (both data-dependent and data-independent) may therefore present additional opportunities for the characterisation of novel exposures and significant research efforts will likely utilise such methods in the future.

Overall, the literature suggests that a paradigm shift is occurring with the coalescence of the analytical approaches used for data acquisition in both untargeted metabolomics and chemical exposomics.

Simultaneous measurement of the metabolome and the chemical exposome: data analysis

Besides the analytical methods, instrumentation and the shared chemical space, the data curation and processing pipelines also have a significant overlap, with multiple available tools for peak picking, deconvolution, and alignment (MS DIAL [85], XCMS [86], MZmine3 [87]) being suitable for both metabolomics and chemical exposomics.

After the primary processing (peak picking, deconvolution, and alignment) of the obtained datasets, it is necessary to apply batch and drift correction [88], filtering, and normalization to the data [89]. At these steps, the sets of well-established guidelines and common practices used in metabolomics [90] may need to be adapted for hybrid metabolome-exposome datasets for several reasons. First, peak detection rate thresholds used in metabolomics studies assume that endogenous features will be ubiquitously detected (usually 50–100 % detection rate [91]). Second, a quality control (QC) filtering step prompts the exclusion of features with coefficients of variation typically higher than 30 % [90]. Both filtering methods would exclude an important fraction of environmental exposures that may only pass a threshold of detection in a minority of samples and might not be detected in the QC. Detection rate thresholds should be relaxed even to detection in low percentages of study samples (e.g., 33 % [14]), if robust associations to the phenotype are found and subsequent level 1 identification can be achieved [92]. Acceptance of features with low detection frequencies encompasses the need for the development of data handling and statistical tools for sparse mass spectrometry data, predominantly imputation algorithms [93,94].

Non-detection of a feature in the study QC also poses additional challenges. Many metabolomics studies have traditionally used QC filtering and only performed MS/MS fragmentation for identification purposes on the QC sample. MS/MS fragmentation in every sample could be prioritized in the chemical exposomics context if fast-scanning instruments are available. The QC filtering step may not be appropriate at all for features with low detection rate (i.e., non-detection in the QC) and could be replaced by adaptive strategies (e.g., dichotomization triggered whenever the detection rate is lower than 50 % [91]). Alternative approaches include the use of separate pooled QC samples for study subpopulations with substantially distinct exposure profiles (e.g., smokers vs. non-smokers [79]), or the injection of a standard mix including different compound classes throughout the analytical run, mostly applicable to suspect screening approaches [95].

The main bottleneck in achieving a more detailed picture of the metabolome and the chemical exposome through HRMS continues to be feature identification. One major limitation is the greatly variable proportion of MS/MS spectra acquired that will not produce an adequate match to a spectral library. This can occur due to a variety of technical reasons: multiple precursors in an isolation width, coelution of isomers, incorrect deconvolution, instrumental noise, or simply the unfeasibility of optimizing collision energy settings for very different compound classes, leading to excessive fragmentation for some species and insufficient fragmentation for others. While the fraction of good quality spectra is likely to increase along technical advances, another reason for the partial identification of HRMS datasets is the lack of libraries that gather comprehensive spectral information of both the endogenous metabolome and the vast (and growing) chemical space humans can be exposed to. Whereas the endogenous human metabolome is relatively well-characterized, with long-established databases (HMDB [83], Lipid Maps [96], METLIN [97] ...), the universe of exogenous chemicals to which humans can be exposed is much more diverse [98]. Although there are exogenous chemicals databases available, often focusing on specific kinds of exposures (e.g., food [99], toxins [100], natural products [101], pesticides (https://food.ec.europa.eu/plants/pesticides/eu-pesticides-database_en), etc.), these do not always contain spectral information. On the other hand, commercial spectral libraries (such as NIST or Wiley MS) can be valuable tools to simultaneously boost identifications of the metabolome and to have an overview of the chemical exposome. These can be used in conjunction with specialized libraries, usually more comprehensive for a given exposure family of interest, or with injected in-house standards that align with the goals of a particular research project.

In essence, there is a compelling need to comprehensively investigate all these forms of exposures through the lens of the chemical exposome concept, and in a more dynamic, exposomics community-driven, and open-source manner. The annotation gap between metabolomics-focused studies and chemical exposome-focused studies is also gradually being bridged with open-source broad-spectrum initiatives and compound libraries available (PubChemLite [102], NORMAN (<https://www.norman-network.com/>), the blood exposome database [103], Exposome Explorer [104], or the MS DIAL open-source spectral library [85]), together with tools such as advanced retention time prediction models [84] or *in silico* MS/MS fragmentation prediction (MetFrag [102], SIRIUS [105]).

Simultaneous measurement of the metabolome and the chemical exposome: data interpretation

At the core of exposome science we find the characterization of the associations between the mixture of measured exposures and health outcomes. These analyses are conducted at the higher order phenotypic level, e.g., on cancer or cardio-metabolic outcomes. To compute exposome–outcome associations advanced statistical tools are required, usually applying feature selection or feature reduction to deal with the high dimensionality of the typical exposome dataset. Several approaches explicitly search for combined effects of exposures and/or their interactions using linear index models or response surface methods, in addition to simply searching for single exposure effect using an exposome-wide association study (ExWAS) [106].

Measuring and reporting exposure–outcome associations is important evidence to support reduction in the environment through regulatory action (e.g., forbidden components of pesticides or PCCPs). However, often such information is insufficient for certain classes of exposures (e.g., natural sources, no industrial alternative, or not enough social consensus for their prohibition). In this context, it becomes necessary to consider the mechanisms or intermediate effects within the human body for effective intervention, risk mitigation or even treatment of adverse effects. In this regard, integration of exposomics data with other omics [13,14,75], though currently at its infancy, holds great potential. Indeed, exposome–health associations at the molecular level (i.e., using omics approaches) may identify more subtle biological responses, before the appearance of clinical symptoms. Strong response signatures at the molecular levels in observational human studies may also support the regulation of chemical exposures through the understanding of toxicological mechanism [107] and the construction of new AOPs [12]. In addition, the characterization and validation of novel molecular mechanisms can be further supported by *in vitro*

and animal studies, or intervention studies that evaluate tissue-specific or circulating biomarkers.

Another partially unsolved fundamental challenge of combining untargeted metabolomics and chemical exposomics relies in distinguishing the subset of features comprising a causal signature (exposures) from those indicating the biomolecular response to the exposures on human health (the metabolome). It is also possible that some molecules can simultaneously fall into both categories (secondary metabolism of exposures by human enzymes and gut microbiota). This was well exemplified in an intervention study assessing the impact of swimming pool disinfection by products on volunteers metabolome [75]. Adding another layer of complexity, there is the need to disentangle the dynamic component of these observations, and this will require longitudinal as well as cross-sectional studies. For instance, performing a completely cross-sectional integration analysis [13], our team recently found that chemical exposures associated with diet-related metabolites, specifically pesticides with the fruit biomarker proline betaine and heavy metals with fish-related PUFAs in children [6]. However, while these associations may be informative about the routes of exposure to chemicals, the additional predictive value of the dietary markers is less clear. Others have applied longitudinal analyses to link the exposome of the first trimester of pregnancy to the metabolome of the third trimester [77], have evaluated the temporal variation of a single subject over an extended period [30] or proposed the daunting task of analysis across a participant's whole lifespan [108]. Nevertheless, the current lack of consensus in study designs and the independent spatiotemporal variability of both the metabolome and the exposome (with confounders such as year of sample collection or seasonality) currently present severe challenges to replication and validation of observations in the field of exposomics.

In addition, the wide variety of chromatographic methods, combinations of MS applications, use of different libraries for compound identification, and reporting standards currently used in untargeted exposome research studies, as exemplified in Table 3, also contribute to the difficulty of replicating findings and harmonization of data across exposomics studies. Whereas data harmonization across projects is an established reality for other types of exposome variables and other omics [109], small-molecule HRMS exposomics data harmonization, despite valuable and significant efforts [110,111], still has a long path to tread.

Conclusions

Determining the chemical exposome will always require a blend of targeted and untargeted HRMS approaches. A continuous cyclic process is needed between untargeted

HRMS to drive the discovery of new exposures with implications for health outcomes; and targeted MS or HRMS to provide rigorous quantitative data for human biomonitoring studies.

Bridging the gap between the metabolome and the chemical exposome is possible due to their substantial chemical overlap. However, achieving this will require the consolidation of compound and spectral libraries that comprehensively cover both domains. Other approaches that could contribute are the acquisition of MS/MS on every sample and the relaxation of data filtering procedures (particularly QC-based filtering and detection rate filtering) to cope with the sparsity intrinsic to chemical exposures. Moreover, a more comprehensive view of the metabolome and the chemical exposome can be achieved if complementary HRMS assays are employed (GC and LC, or multiple LC assays with complementary column chemistries), as exemplified throughout this review. Additionally, a more defined framework for categorizing compounds into endogenous or exogenous categories, or exogenous origins that involve human or microbial metabolism may also be of benefit.

Finally, while we have described how the simultaneous metabolome and chemical exposome characterization can be achieved and facilitated, it remains a challenge to interpret these data in an integrated way, due to the differences in how the metabolome and exposome vary over time and the timescales of resulting biological effects. It is clear that more longitudinal studies with both metabolome and exposome data available will be needed to fully capture and unravel the effect of exposures on human health.

Author contributions

CB: Conceptualization and writing of the original draft, data curation and visualization, review and editing. YX: Data curation and visualization, writing - review and editing. RGS, LM, PGF: Writing of the original draft, review and editing. HCK: Conceptualization, writing of the original draft, review and editing.

Editorial disclosure statement

Hector Keun had no involvement in the peer review of the article and has no access to information regarding its peer-review. Full responsibility for the editorial process of this article was delegated to James McCullagh.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

CB, LM and HCK received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 874583 [ATHLETE]. YX is supported by the President's PhD Scholarship from Imperial College London.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest

1. Wild CP: **Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology.** *Cancer Epidemiol Biomark Prev* 2005, **14**:1847–1850.
- Wild introduces the "exposome" concept for the first time in this opinion paper.
2. Miller GW, Jones DP: **The nature of nurture: refining the definition of the exposome.** *Toxicol Sci* 2014, **137**:1–2.
3. Price EJ, Vitale CM, Miller GW, David A, Barouki R, Audouze K, Walker DI, Antignac J-P, Coumoul X, Bessonneau V, *et al.*: **Merging the exposome into an integrated framework for omics sciences.** 2022, <https://doi.org/10.1016/j.isci>.
4. Gil-Solsona R, Nika MC, Bustamante M, Villanueva CM, Foraster M, Cosin-Tomás M, Alygizakis N, Gómez-Roig MD, Llurba-Olive E, Sunyer J, *et al.*: **The potential of sewage sludge to predict and evaluate the human chemical exposome.** *Environ Sci Technol Lett* 2021, **8**:1077–1084.
- Finding efficient proxies to predict the risk of chemical exposure in humans is an urgent need to cover large areas and populations at a reasonable cost. This paper provides interesting data on the occurrence of contaminants in blood and placenta (using non-target approaches) and propose the use of sewage epidemiology (or wastewater based epidemiology) as a complementary tool to evaluate the exposome in large populations without the need of conducting biomonitoring campaigns.
5. Goin DE, Abrahamsson D, Wang M, Jiang T, Park JS, Sirota M, Morello-Frosch R, DeMicco E, Zlatnik MG, Woodruff TJ: **Disparities in chemical exposures among pregnant women and neonates by socioeconomic and demographic characteristics: a nontargeted approach.** *Environ Res* 2022:215.
6. Rojas-Rueda D, Nieuwenhuijsen MJ, Gascon M, Perez-Leon D, Mudu P: **Green spaces and mortality: a systematic review and meta-analysis of cohort studies.** *Lancet Planet Health* 2019, **3**:e469–e477.
7. Nutsford D, Pearson AL, Kingham S, Reitsma F: **Residential exposure to visible blue space (but not green space) associated with lower psychological distress in a capital city.** *Health Place* 2016, **39**:70–78.
8. Ikeda E, Hinckson E, Witten K, Smith M: **Associations of children's active school travel with perceptions of the physical environment and characteristics of the social environment: a systematic review.** *Health Place* 2018, **54**:118–131.
9. Vermeulen R, Schymanski EL, Barabási AL, Miller GW: **The exposome and health: where chemistry meets biology.** *Science* 2020:392–396. 1979.
- This review discusses progress in assessing the chemical component of the exposome and its implications for human health. The authors delineate the roles of HRMS and network sciences on taking the first steps toward a comprehensive assessment of the exposome. The authors discuss the necessity to characterize the exposome at a scale comparable to the current understanding of the human genome.
10. Niedzwiecki MM, Walker DI, Vermeulen R, Chadeau-Hyam M, Jones DP, Miller GW: **The exposome: molecules to populations.** *Annu Rev Pharmacol Toxicol* 2019, **59**:107–127.

11. Escher BI, Hackermüller J, Polte T, Scholz S, Aigner A, Altenburger R, Böhme A, Bopp SK, Brack W, Busch W, *et al.*: **From the exposome to mechanistic understanding of chemical-induced adverse effects.** *Environ Int* 2017, **99**:97–106.
12. Knapen D, Angrish MM, Fortin MC, Katsiadaki I, Leonard M, Margiotta-Casaluci L, Munn S, O'Brien JM, Pollesch N, Smith LC, *et al.*: **Adverse outcome pathway networks I: development and applications.** *Environ Toxicol Chem* 2018, **37**:1723–1733.
13. Maitre L, Bustamante M, Hernández-Ferrer C, Thiel D, Lau CHE, Siskos AP, Vives-Usano M, Ruiz-Arenas C, Pelegrí-Sisó D, Robinson O, *et al.*: **Multi-omics signatures of the human early life exposome.** *Nat Commun* 2022, **13**.
- A unique example of integration of 100s of exposures with multi-omics datasets (epigenome, transcriptome, proteins and metabolites) in 1300 mother-child pairs. They identified novel signatures of chemical and physical exposures on the metabolome of children, identified novel pathways of exposures and responses.
14. Gao P, Shen X, Zhang X, Jiang C, Zhang S, Zhou X, Schüssler-Florenz Rose SM, Snyder M: **Precision environmental health monitoring by longitudinal exposome and multi-omics profiling.** *Genome Res* 2022, **32**:1199–1214.
- Longitudinal multiomics-exposome integration study at the level of the individual. They highlight multiple associations between the personal external exposome and immune, renal, and hepatic function.
15. Juarez PD, Hood DB, Rogers GL, Baktash SH, Saxton AM, Matthews-Juarez P, Im W, Cifuentes MP, Phillips CA, Lichtveld MY, *et al.*: **A novel approach to analyzing lung cancer mortality disparities: using the exposome and a graph-theoretical toolchain HHS Public Access.** *Environ Dis* 2017, **2**:33–44.
16. Polemi KM, Nguyen VK, Heidt J, Kahana A, Jolliet O, Colacino JA: **Identifying the link between chemical exposures and breast cancer in African American women via integrated in vitro and exposure biomarker data.** *Toxicology* 2021:463.
17. Ronsmans S, Sørig Hougaard K, Nawrot TS, Plusquin M, Huaux F, Jesús Cruz M, Moldovan H, Verpaele S, Jayapala M, Tunney M, *et al.*: **The EXIMIOUS project-Mapping exposure-induced immune effects: connecting the exposome and the immunome.** *Environ Epidemiol* 2022, **6**:E193.
18. Laiho JE, Laitinen OH, Malkamäki J, Puustinen L, Sinkkonen A, Pärkkä J, Hyöty H, Anna E, Heikki H, Kalle K, *et al.*: **Exposomic determinants of immune-mediated diseases: special focus on type 1 diabetes, celiac disease, asthma, and allergies: the HEDIMED project approach.** *Environ Epidemiol* 2022, **6**:E212.
19. Vlaanderen J, De Hoogh K, Hoek G, Peters A, Probst-Hensch N, Scalbert A, Melén E, Tonne C, De Wit GA, Chadeau-Hyam M, *et al.*: **Developing the building blocks to elucidate the impact of the urban exposome on cardiometabolic-pulmonary disease: the EU EXPANSE project.** *Environ Epidemiol* 2021:5.
20. Stratakis N, Conti DV, Jin R, Margetaki K, Valvi D, Siskos AP, Maitre L, Garcia E, Varo N, Zhao Y, *et al.*: **Prenatal exposure to perfluoroalkyl substances associated with increased susceptibility to liver injury in children HHS public access.** *Hepatology* 2020, **72**:1758–1770.
21. Paul KC, Krolewski RC, Lucumi Moreno E, Blank J, Holton KM, Ahfeldt T, Furlong M, Yu Y, Cockburn M, Thompson LK, *et al.*: **A pesticide and iPSC dopaminergic neuron screen identifies and classifies Parkinson-relevant pesticides.** *Nat Commun* 2023:14.
22. Ladvá CN, Golan R, Liang D, Greenwald R, Walker DI, Uppal K, Raysoni AU, Tran VL, Yu T, Dana Flanders W, *et al.*: **Particulate metal exposures induce plasma metabolome changes in a commuter panel study.** *PLoS One* 2018, **13**.
23. Pronk A, Loh M, Kuijpers E, Albin M, Selander J, Godderis L, Ghosh M, Vermeulen R, Peters S, Mehlum IS, *et al.*: **Applying the exposome concept to working life health: the EU EPHOR project.** *Environ Epidemiol* 2022, **6**:E185.
24. Rotander A, Kärman A, Toms LML, Kay M, Mueller JF, Gómez Ramos MJ: **Novel fluorinated surfactants tentatively identified in firefighters using liquid chromatography quadrupole time-**

- of-flight tandem mass spectrometry and a case-control approach.** *Environ Sci Technol* 2015, **49**:2434–2442.
25. Vrijheid M, Slama R, Robinson O, Chatzi L, Coen M, van den Hazel P, Thomsen C, Wright J, Athersuch TJ, Avellana N, et al.: **The human early-life exposome (HELIX): project rationale and design.** *Environ Health Perspect* 2014, **122**:535–544.
 26. Balshaw DM, Collman GW, Gray KA, Thompson CL: **The Children's Health Exposure Analysis Resource: enabling research into the environmental influences on children's health outcomes.** *Curr Opin Pediatr* 2017, **29**:385–389.
 27. Panagopoulos Abrahamsson D, Wang A, Jiang T, Wang M, Siddharth A, Morello-Frosch R, Park JS, Sirota M, Woodruff TJ: **A comprehensive non-targeted analysis study of the prenatal exposome.** *Environ Sci Technol* 2021, **55**:10542–10557.
 28. Xia X, Zheng Y, Tang X, Zhao N, Wang B, Lin H, Lin Y: **Nontarget identification of novel per- and polyfluoroalkyl substances in cord blood samples.** *Environ Sci Technol* 2022, **56**:17061–17069.
- This study revealed the presence of several novel PFASs in cord blood samples for the first time and evaluated the exposure risks of identified PFASs.
29. Duszka HM, Manz KE, Pennell KD, Kanda R, Legler J: **Identification of known and novel nonpolar endocrine disruptors in human amniotic fluid.** *Environ Int* 2022, **158**, 106904.
- This study reveals the presence of a wide variety of nonpolar EDCs in direct fetal environment and for the first time identifies novel EDCs in human amniotic fluid. Further assessment of the source and extent of human fetal exposure to these compounds is warranted.
30. Braun D, Schernhammer E, Marko D, Warth B: **Longitudinal assessment of mycotoxin co-exposures in exclusively breastfed infants.** *Environ Int* 2020, **142**, 105845.
 31. Jamnik T, Flasch M, Braun D, Fareed Y, Wasinger D, Seki D, Berry D, Berger A, Wisgrill L, Warth B: **Next-generation biomonitoring of the early-life chemical exposome in neonatal and infant development.** *Nat Commun* 2022, **13**:1–14.
- A MRM method for the determination of multiple exposure families (N = 81) in multiple matrices is presented, including the detection of 27 xenobiotics (plasticizers, PFAS, parabens) in plasma of extremely-premature infants and 29 analytes detected in breast milk. Also, a preliminary estimation of daily toxicant intake via breast milk is conducted.
32. Lu D, Wang D, Ni R, Lin Y, Feng C, Xu Q, Jia X, Wang G, Zhou Z: **Organochlorine pesticides and their metabolites in human breast milk from Shanghai, China.** *Environ Sci Pollut Control Ser* 2015, **22**:9293–9306.
 33. Croes K, Colles A, Koppen G, De Galan S, Vandermarken T, Govarts E, Bruckers L, Nelen V, Schoeters G, Van Larebeke N, et al.: **Determination of PCDD/Fs, PBDD/Fs and dioxin-like PCBs in human milk from mothers residing in the rural areas in Flanders, using the CALUX bioassay and GC-HRMS.** *Talanta* 2013, **113**:99–105.
 34. Athersuch TJ, Keun HC: **Metabolic profiling in human exposure studies.** *Mutagenesis* 2015, **30**:755–762.
 35. Yang T, Hui R, Nouws J, Sauler M, Zeng T, Wu Q: **Untargeted metabolomics analysis of esophageal squamous cell cancer progression.** *J Transl Med* 2022:20.
 36. Szczerbinski L, Wojciechowska G, Olichwier A, Taylor MA, Puchta U, Konopka P, Paszko A, Citek A, Gosciak J, Fiehn O, et al.: **Untargeted metabolomics analysis of the serum metabolic signature of childhood obesity.** *Nutrients* 2022, **14**.
 37. Proitsi P, Kim M, Whitley L, Simmons A, Sattlecker M, Velayudhan L, Lupton MK, Soininen H, Kloszewska I, Mecocci P, et al.: **Association of blood lipids with Alzheimer's disease: a comprehensive lipidomics analysis.** *Alzheimer's Dementia* 2017, **13**:140–151.
 38. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, Sinelnikov I, Krishnamurthy R, Eisner R, Gautam B, et al.: **The human serum metabolome.** *PLoS One* 2011, **6**, e16957.
 39. Bouatra S, Aziat F, Mandal R, Guo AC, Wilson MR, Knox C, Bjorn Dahl TC, Krishnamurthy R, Saleem F, Liu P, et al.: **The human urine metabolome.** *PLoS One* 2013, **8**, e73076.
 40. CDC: **National Health and Nutrition Examination Survey (NHANES 2011–2016).** 2019. <https://www.cdc.gov/nchs/nhanes/index.htm>.
 41. HBM4EU: **The human biomonitoring for the European union network -**. 2022. <https://www.hbm4eu.eu/>.
 42. Contrepas K, Jiang L, Snyder M: **Optimized analytical procedures for the untargeted metabolomic profiling of human urine and plasma by combining hydrophilic interaction (HILIC) and reverse-phase liquid chromatography (RPLC)-mass spectrometry.** *Mol Cell Proteomics* 2015, **14**:1684–1695.
 43. González-Domínguez R, Jáuregui O, Queipo-Ortuño MI, Andrés-Lacueva C: **Characterization of the human exposome by a comprehensive and quantitative large-scale multianalyte metabolomics platform.** *Anal Chem* 2020, **92**:13767–13775.
- A MRM triple-quadrupole targeted method for the determination of >700 analytes comprising human polar metabolites, food exposures, antibiotics and xenobiotics.
44. Silva MJ, Wong LY, Samandar E, Preau JL, Calafat AM, Ye X: **Exposure to di-2-ethylhexyl terephthalate in a convenience sample of U.S. adults from 2000 to 2016.** *Arch Toxicol* 2017, **91**:3287–3291.
 45. Prasain JK, Arabshahi A, Moore DR, Greendale GA, Wyss JM, Barnes S: **Simultaneous determination of 11 phytoestrogens in human serum using a 2 min liquid chromatography/tandem mass spectrometry method.** *J Chromatogr, B: Anal Technol Biomed Life Sci* 2010, **878**:994–1002.
 46. Wang Y, Meng L, Pittman EN, Etheredge A, Hubbard K, Trinidad DA, Kato K, Ye X, Calafat AM: **Quantification of urinary mono-hydroxylated metabolites of polycyclic aromatic hydrocarbons by on-line solid phase extraction-high performance liquid chromatography-tandem mass spectrometry.** *Anal Bioanal Chem* 2017, **409**:931–937.
 47. Xia B, Xia Y, Wong J, Nicodemus KJ, Xu M, Lee J, Guillot T, Li J: **Quantitative analysis of five tobacco-specific N-nitrosamines in urine by liquid chromatography-atmospheric pressure ionization tandem mass spectrometry.** *Biomed Chromatogr* 2014, **28**:375–384.
 48. Bernert JT, McGuffey JE, Morrison MA, Pirkle JL: **Comparison of serum and salivary cotinine measurements by a sensitive high-performance liquid chromatography-tandem mass spectrometry method as an indicator of exposure to tobacco smoke among smokers and nonsmokers.** *J Anal Toxicol* 2000, **24**:333–339.
 49. Zhang L, Xia Y, Xia B, Nicodemus KJ, McGuffey J, McGahee E, Blount B, Wang L: **High-throughput and sensitive analysis of urinary heterocyclic aromatic amines using isotope-dilution liquid chromatography-tandem mass spectrometry and robotic sample preparation system.** *Anal Bioanal Chem* 2016, **408**:8149–8161.
 50. Davis MD, Wade EL, Restrepo PR, Roman-Esteva W, Bravo R, Kuklenyik P, Calafat AM: **Semi-automated solid phase extraction method for the mass spectrometric quantification of 12 specific metabolites of organophosphorus pesticides, synthetic pyrethroids, and select herbicides in human urine.** *J Chromatogr, B: Anal Technol Biomed Life Sci* 2013, **929**:18–26.
 51. Cequier E, Sakhi AK, Haug LS, Thomsen C: **Development of an ion-pair liquid chromatography-high resolution mass spectrometry method for determination of organophosphate pesticide metabolites in large-scale biomonitoring studies.** *J Chromatogr A* 2016, **1454**:32–41.
 52. Olsson AO, Baker SE, Nguyen JV, Romanoff LC, Udunka SO, Walker RD, Flemmen KL, Barr DB: **A liquid chromatography-tandem mass spectrometry multiresidue method for quantification of specific metabolites of organophosphorus pesticides, synthetic pyrethroids, selected herbicides, and DEET in human urine.** *Anal Chem* 2004, **76**:2453–2461.
 53. Mouskeftara T, Virgiliou C, Iakovakis A, Raikos N, Gika HG: **Liquid chromatography tandem mass spectrometry for the determination of nine insecticides and fungicides in human postmortem blood and urine.** *J Chromatogr, B: Anal Technol Biomed Life Sci* 2021:1179.
 54. Kuklenyik P, Baker SE, Bishop AM, Morales -AP, Calafat AM: **On-line solid phase extraction-high performance liquid**

- chromatography-isotope dilution-tandem mass spectrometry approach to quantify N,N-diethyl-m-toluamide and oxidative metabolites in urine. *Anal Chim Acta* 2013, **787**:267–273.
55. Kuklenyik Z, Panuwet P, Jayatilaka NK, Pirkle JL, Calafat AM: **Two-dimensional high performance liquid chromatography separation and tandem mass spectrometry detection of atrazine and its metabolic and hydrolysis products in urine.** *J Chromatogr, B: Anal Technol Biomed Life Sci* 2012, **901**:1–8.
 56. Nguyen JV, Olsson AO, Bravo R, Needham LT, Barr DB: **Quantification of atrazine, phenylurea, and sulfonyleurea herbicide metabolites in urine by high-performance liquid chromatography-tandem mass spectrometry.** *J Anal Toxicol* 2007, **31**:181–186.
 57. Jayatilaka NK, Restrepo P, Williams LT, Ospina M, Valentin-Blasini L, Calafat AM: **Quantification of three chlorinated dialkyl phosphates, diphenyl phosphate, 2,3,4,5-tetrabromobenzoic acid, and four other organophosphates in human urine by solid phase extraction-high performance liquid chromatography-tandem mass spectrometry.** *Anal Bioanal Chem* 2017, **409**:1323–1332.
 58. Sjödin A, Jones RS, Lapeza CR, Focant J-FO, McGahee EE, Patterson DG: **Semiautomated high-throughput extraction and cleanup method for the measurement of polybrominated diphenyl ethers, polybrominated biphenyls, and polychlorinated biphenyls in human serum.** *Anal Chem* 2004, **76**:1921–1927.
 59. Kato K, Kalathil AA, Patel AM, Ye X, Calafat AM: **Per- and polyfluoroalkyl substances and fluorinated alternatives in urine and serum by on-line solid phase extraction-liquid chromatography-tandem mass spectrometry.** *Chemosphere* 2018, **209**:338–345.
 60. Li Y, Xu S, Li Y, Zhang B, Huo W, Zhu Y, Wan Y, Zheng T, Zhou A, Chen Z, et al.: **Association between urinary parabens and gestational diabetes mellitus across prepregnancy body mass index categories.** *Environ Res* 2019, **170**:151–159.
 61. Cao S, Liu Z, Zhang L, Xi C, Li X, Wang G, Yuan R, Mu Z: **Development of an HPLC-MS/MS method for the simultaneous analysis of six kinds of parabens in food.** *Anal Methods* 2013, **5**:1016–1023.
 62. Zhou X, Kramer JP, Calafat AM, Ye X: **Automated on-line column-switching high performance liquid chromatography isotope dilution tandem mass spectrometry method for the quantification of bisphenol A, bisphenol F, bisphenol S, and 11 other phenols in urine.** *J Chromatogr, B: Anal Technol Biomed Life Sci* 2014, **944**:152–156.
 63. Alwis KU, Blount BC, Silva LK, Smith MM, Loose KH: **Method for quantifying nitromethane in blood as a potential biomarker of halonitromethane exposure.** *Environ Sci Technol* 2008, **42**:2522–2527.
 64. Alwis KU, Bailey TL, Patel D, Wang L, Blount BC: **Measuring urinary N-acetyl-S-(4-hydroxy-2-methyl-2-buten-1-yl)-L-cysteine (IPMA3) as a potential biomarker of isoprene exposure.** *Anal Chim Acta* 2016, **941**:61–66.
 65. Zheng X, Dupuis KT, Aly NA, Zhou Y, Smith FB, Tang K, Smith RD, Baker ES: **Utilizing ion mobility spectrometry and mass spectrometry for the analysis of polycyclic aromatic hydrocarbons, polychlorinated biphenyls, polybrominated diphenyl ethers and their metabolites.** *Anal Chim Acta* 2018, **1037**:265–273.
 66. Jursa S, Chovancová J, Petřík J, Lokša J: **Dioxin-like and non-dioxin-like PCBs in human serum of Slovak population.** *Chemosphere* 2006, **64**:686–691.
 67. Zhang L, Zhong Y, Liu X, Bao Y, Zhao Y, Wu Y, Cai Z, Li J: **Determination of polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans, and dioxin-like polychlorinated biphenyls in human serum using programmable-temperature vaporization gas chromatography with high-resolution mass spectrometry.** *J Separ Sci* 2017, **40**:3453–3461.
 68. Hung CC, Sari Simaremare SR, Hsieh CJ, Yiin LM: **Simultaneous determination of pyrethroid, organophosphate and carbamate metabolites in human urine by Gas Chromatography-Mass Spectrometry (GCMS).** *Appl Sci* 2019, **9**.
 69. Aly NA, Dodds JN, Luo YS, Grimm FA, Foster MK, Rusyn I, Baker ES: **Utilizing ion mobility spectrometry-mass spectrometry for the characterization and detection of persistent organic pollutants and their metabolites.** *Anal Bioanal Chem* 2022, **414**:1245–1258.
 70. Flasch M, Fitz V, Rampler E, Ezekiel CN, Koellensperger G, Warth B: **Integrated exposomics/metabolomics for rapid exposure and effect analyses.** *JACS Au* 2022, **2**:2548–2560.
 71. Roggeman M, Belova L, Fernández SF, Kim DH, Jeong Y, Poma G, Remy S, Verheyen VJ, Schoeters G, van Nuijs ALN, et al.: **Comprehensive suspect screening for the identification of contaminants of emerging concern in urine of Flemish adolescents by liquid chromatography high-resolution mass spectrometry.** *Environ Res* 2022:214.
 72. Coscollà C, León N, Pastor A, Yusà V: **Combined target and post-run target strategy for a comprehensive analysis of pesticides in ambient air using liquid chromatography-Orbitrap high resolution mass spectrometry.** *J Chromatogr A* 2014, **1368**:132–142.
 73. Jin R, McConnell R, Catherine C, Xu S, Walker DI, Stratakis N, Jones DP, Miller GW, Peng C, Conti DV, et al.: **Perfluoroalkyl substances and severity of nonalcoholic fatty liver in Children: an untargeted metabolomics approach.** *Environ Int* 2020:134.
 74. Walker DI, Marder ME, Yano Y, Terrell M, Liang Y, Barr DB, Miller GW, Jones DP, Marcus M, Pennell KD: **Multigenerational metabolic profiling in the Michigan PBB registry.** *Environ Res* 2019, **172**:182–193.
 75. van Veldhoven K, Keski-Rahkonen P, Barupal DK, Villanueva CM, Font-Ribera L, Scalbert A, Bodinier B, Grimalt JO, Zwiener C, Vlaanderen J, et al.: **Effects of exposure to water disinfection by-products in a swimming pool: a metabolome-wide association study.** *Environ Int* 2018, **111**:60–70.
- Panel study evaluating the effect of swimming pool disinfection by-products. They measured the endogenous metabolites and exogenous chemicals through LC-HRMS and reporting associations exposome-metabolome associations through a metabolome-wide association study (MWAS).
76. Walker DI, Juran BD, Cheung AC, Schlicht EM, Liang Y, Niedzwiecki M, Larusso NF, Gores GJ, Jones DP, Miller GW, et al.: **High-resolution exposomics and metabolomics reveals specific associations in cholestatic liver diseases.** *Hepatol Commun* 2021, **6**:2022.
 77. Chen M, Guan Y, Huang R, Duan J, Zhou J, Chen T, Wang X, Xia Y, London SJ: **Associations between the maternal exposome and metabolome during pregnancy.** *Environ Health Perspect* 2022:130.
- Chen et al. measured the metabolome and the chemical exposome in >1000 pregnant women in the first and third trimesters of pregnancy using a common LC-HRMS method. They identified around a thousand associations between chemical exposures and maternal metabolism during pregnancy.
78. Zhang P, Carlsten C, Chaleckis R, Hanhineva K, Huang M, Isobe T, Koistinen VM, Meister I, Papazian S, Sdougkou K, et al.: **Defining the scope of exposome studies and research needs from a multidisciplinary perspective.** *Environ Sci Technol Lett* 2021, **8**:839–852.
 79. Frigerio G, Moruzzi C, Mercadante R, Schymanski EL, Fustinoni S: **Development and application of an LC-MS/MS untargeted exposomics method with a separated pooled quality control strategy.** *Molecules* 2022, **27**.
 80. Li YY, Rushing B, Schroder M, Sumner S, Kay CD: **Exploring the contribution of (Poly)phenols to the dietary exposome using high resolution mass spectrometry untargeted metabolomics.** *Mol Nutr Food Res* 2022:66.
 81. Rusli H, Putri RM, Alni A: **Recent developments of liquid chromatography stationary phases for compound separation: from proteins to small organic compounds.** *Molecules* 2022, **27**.

82. Sdougkou K, Xie H, Papazian S, Bonnefille B, Bergdahl IA, Martin JW: **Phospholipid removal for enhanced chemical exposomics in human plasma.** *Environ Sci Technol* 2023, <https://doi.org/10.1021/acs.est.3c00663>.
- Chemical exposomics in human plasma was enhanced by an optimized phospholipid removal step that increased targeted method sensitivity while also revealing >13,000 new molecular features by LC-HRMS non-targeted acquisition.
83. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Jewell K, Arndt D, Sawhney S, *et al.*: **HMDB: the human metabolome database.** *Nucleic Acids Res* 2007, **35**:521–526.
84. Bonini P, Kind T, Tsubawa H, Barupal DK, Fiehn O: **Retip: retention time prediction for compound annotation in untargeted metabolomics.** *Anal Chem* 2020, **92**:7515–7522.
85. Tsubawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, Vanderghenst J, Fiehn O, Arita M: **MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis.** *Nat Methods* 2015, **12**:523–526.
86. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G: **XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification.** *Anal Chem* 2006, **78**:779–787.
87. Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrland TS, Bushuev R, Murray KJ, Hoffmann N, Lu M, *et al.*: **Integrative analysis of multimodal mass spectrometry data in MZmine 3.** *Nat Biotechnol* 2023, **41**:447–449.
88. Brunius C: **Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction.** *Metabolomics* 2016, **12**:1–13.
89. Wu Y, Li L: **Sample normalization methods in quantitative metabolomics.** *J Chromatogr A* 2016, **1430**:80–95.
90. Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, Lewis MR, Dunn WB: **Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies.** *Metabolomics* 2018, **14**:1–17.
- This paper provides extensively followed guidelines in the metabolomics community for data acquisition and QA/QC strategy.
91. Liu RX, Koyawala N, Thiessen-Philbrook HR, Doshi MD, Reese PP, Hall IE, Mohan S, Parikh CR: **Untargeted metabolomics of perfusate and their association with hypothermic machine perfusion and allograft failure.** *Kidney Int* 2023, **103**:762–771.
92. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J: **Identifying small molecules via high resolution mass spectrometry: communicating confidence.** *Environ Sci Technol* 2014, **48**:2097–2098.
93. Wei R, Wang J, Su M, Jia E, Chen S, Chen T, Ni Y: **Missing value imputation approach for mass spectrometry-based metabolomics data.** *Sci Rep* 2018, **8**.
94. Do KT, Wahl S, Raffler J, Molnos S, Laimighofer M, Adamski J, Suhre K, Strauch K, Peters A, Gieger C, *et al.*: **Characterization of missing values in untargeted MS-based metabolomics data and evaluation of missing data handling strategies.** *Metabolomics* 2018, **14**.
95. Acar E, Gürdeniz G, Khakimov B, Savorani F, Korndal SK, Larsen TM, Engelsen SB, Astrup A, Dragsted LO: **Biomarkers of individual foods, and separation of diets using untargeted LC-MS-based plasma metabolomics in a randomized controlled trial.** *Mol Nutr Food Res* 2019, **63**.
96. Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK, Merrill AH, Murphy RC, Raetz CRH, Russell DW, *et al.*: **LMSP: LIPID MAPS structure database.** *Nucleic Acids Res* 2007, **35**:527–532.
97. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G: **METLIN A metabolite mass spectral database.** [date unknown].
98. Gago-Ferrero P, Krettek A, Fischer S, Wiberg K, Ahrens L: **Suspect screening and regulatory databases: a powerful combination to identify emerging micropollutants.** *Environ Sci Technol* 2018, **52**:6881–6894.
99. Harrington RA, Adhikari V, Rayner M, Scarborough P: **Nutrient composition databases in the age of big data: FoodDB, a comprehensive, real-time database infrastructure.** *BMJ Open* 2019, **9**:1–10.
100. Wishart D, Arndt D, Pon A, Sajed T, Guo AC, Djoumbou Y, Knox C, Wilson M, Liang Y, Grant J, *et al.*: **T3DB: the toxic exposome database.** *Nucleic Acids Res* 2015, **43**:D928–D934.
101. Sorokina M, Steinbeck C: **Review on natural products databases: where to find data in 2020.** *J Cheminf* 2020:12.
102. Schymanski EL, Kondić T, Neumann S, Thiessen PA, Zhang J, Bolton EE: **Empowering large chemical knowledge bases for exposomics: PubChemLite meets MetFrag.** *J Cheminf* 2021, **13**:1–15.
103. Barupal DK, Fiehn O: **Generating the blood exposome database using a comprehensive text mining and database fusion approach.** *Environ Health Perspect* 2019, **127**:1–10.
104. Neveu V, Moussy A, Rouaix H, Wedekind R, Pon A, Knox C, Wishart DS, Scalbert A: **Exposome-Explorer: A manually-curated database on biomarkers of exposure to dietary and environmental factors.** *Nucleic Acids Res* 2017, **45**:D979–D984.
105. Dührkop K, Fleischauer M, Ludwig M, Aksenov AA, Melnik AV, Meusel M, Dorrestein PC, Rousu J, Böcker S: **SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information.** *Nat Methods* 2019, **16**:299–302.
106. Maitre L, Guimbaud JB, Warembourg C, Güil-Oumrait N, Petrone PM, Chadeau-Hyam M, Vrijheid M, Basagaña X, Gonzalez JR, Alfano R, *et al.*: **State-of-the-art methods for exposure-health studies: results from the exposome data challenge event.** *Environ Int* 2022:168.
107. Barouki R, Audouze K, Becker C, Blaha L, Coumoul X, Karakitsios S, Klanova J, Miller GW, Price EJ, Sarigiannis D: **The exposome and toxicology: a win-win collaboration.** *Toxicol Sci* 2022, **186**:1–11.
108. Kalia V, Belsky DW, Baccarelli AA, Miller GW: **An exposomic framework to uncover environmental drivers of aging.** *Exposome* 2022, **2**.
- This paper provides an overview of approaches and challenges in using an exposomic framework to investigate environmental drivers of aging. While capturing exposures over a life course is a daunting and expensive task, the use of historical data can be a practical way to approach this research.
109. Schmidt RA, Wey TW, Harding KD, Fortier I, Atkinson S, Tough S, Letourneau N, Knight JA, Fraser WD, Bocking A: **A harmonized analysis of five Canadian pregnancy cohort studies: exploring the characteristics and pregnancy outcomes associated with prenatal alcohol exposure.** *BMC Pregnancy Childbirth* 2023:23.
110. Vineis P, Chadeau-Hyam M, Gmuender H, Gulliver J, Herceg Z, Kleinjans J, Kogevinas M, Kyrtopoulos S, Nieuwenhuijsen M, Phillips DH, *et al.*: **The exposome in practice: design of the EXPOsOMICS project.** *Int J Hyg Environ Health* 2017, **220**:142–151.
111. Ulrich EM, Sobus JR, Grulke CM, Richard AM, Newton SR, Strynar MJ, Mansouri K, Williams AJ: **EPA's non-targeted analysis collaborative trial (ENTACT): genesis, design, and initial findings.** *Anal Bioanal Chem* 2019, **411**:853–866.