Analytical aspects of meet-in-metabolite analysis for molecular pathway reconstitution from exposure to adverse outcome

Heqing Shen[1,*], Yike Zhang[1], Karl-Werner Schramm[2,*]

[1]State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, School of Public Health, Xiamen University, 361102, Xiamen, PR China
[2]Helmholtz Zentrum München, Molecular EXposomics, German Research Center for Environmental Health, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

*Correspondent: hqshen@xmu.edu.cn; schramm@helmholtz-muenchen.de

Abstract: To explore the etiology of diseases is one of the major goals in epidemiological study. Meet-in-metabolite analysis reconstitutes biomonitoring-based adverse outcome (AO) pathways from environmental exposure to a disease, in which the chemical exposome-related metabolism responses are transmitted to incur the AO-related metabolism phenotypes. However, the ongoing data-dependent acquisition of non-targeted biomonitoring by high-resolution mass spectrometry (HRMS) is biased against the low abundance molecules, which forms the major of molecular internal exposome, i.e., the totality of trace levels of environmental pollutants and/or their metabolites in human samples. The recent development of data-independent acquisition protocols for HRMS screening has opened new opportunities to enhance unbiased measurement of the extremely low abundance molecules, which can encompass a wide range of analytes and has been applied in metabolomics, DNA, and protein adductomics. In addition, computational MS for small molecules is urgently required for the top-down exposome databases. Although a holistic analysis of the exposome and endogenous metabolites is plausible, multiple and flexible strategies, instead of "putting one thing above all" are proposed.

Keywords: human biomonitoring; molecular exposome, metabolome, non-targeted analysis; adverse outcome pathway; system epidemiology

Introduction

Although it is considered a "black box" epidemiology, the traditional epidemiology founded the identification of disease etiologies (Helzlsouer, 1993), where an exposure factor is linked to a disease through statistical association. From exposure to prognostic significance, molecular epidemiology has incorporated a series of biomarkers; these biomarkers can be scattered and suggest pathways that may lead to disease initiation or development based on risk factor occurrence. Successful cases are rare if all the biomarkers are available along the exposure-adverse outcome course. These biomarkers have been individually investigated, and their assessments focus on the reliability and validity of statistical models such as receiver operating characteristic- or sensitive curve-based models. Biological relationships between the upstream and

downstream biomarkers, also known as biomarker webs, are typically omitted. Therefore, molecular epidemiology is contingent on how well the data reflect the events and involve the risk factor that induced the AO pathways (AOPs). One well-known example in molecular epidemiology is human exposure to benzo[a]pyrene (BP) and the resulting cancer risk. The glucuronide conjugate of BP-7,8-dihydrodiol in urine can be measured as the exposure biomarker. The further metabolism of BP-7,8-dihydrodiol-9,10-epoxide combination with DNA bases, repairing mutation initiation, and accumulation of urinary 8-hydroxy-2-deoxyguanosine can mark an adverse effect and may cause a cancer diagnosis. However, these biomarkers alone cannot profile a pathway to determine the extent to which environmental risk leads to disease development. Systems biology and molecular epidemiology have combined to introduce the systems epidemiology era. The details of molecular events in response to both environmental exposure and disease or adverse health outcomes can be profiled. These molecular events form a biological continuum on the data of genomics, proteomics, and metabolomics. High-throughput data acquisition has been sufficiently applied for omic-wide associated studies, including effect-oriented exposome-wide and/or metabolome-wide studies. Although exposure and outcome molecular biomarkers have been more recently screened using a holistic view instead of a reductive one, the pathway reconstitution on systems biology, i.e., uncovering the black box between the exposure and AO, has not yet been fully considered. In this review, we firstly profiled the ongoing development of meet-in-metabolite analysis (MIMA) for the AOPs' reconstitution on metabolism (Wu et al., 2021), and then discussed the analytical choke-point of low level molecules in exposomics and metabolomics. Because the latter would hinder the quality of these multiple omics database in the MIMA procedure, the advances, techniques, developments and outlooks of the unbiased traceability of total internal exposome (i.e., chemical pollutants and metabolites) were majorly reviewed, which are important and forms the foundation of the suggested systems epidemiological paradigm.

Metabolism-based pathway reconstitution from exposure to disease

Traditional human biomonitoring only focuses on the environmental pollutant measurement in populations, which can include parental chemicals and/or their metabolites. The holistic measurement of endogenous metabolites in humans is typically considered in the context of metabolomics. Non-targeted measurements have involved both endogenous and exogenous small molecules in humans. According to the molecular epidemiology framework, xenobiotic biomonitoring data can relate to metabolic response data. Similarly, a diagnostic outcome may be a reflection of certain metabolic phenotypes. Therefore, it is reasonable to merge the two metabolic modules (endogenous and exogenous) on the metabolome continuum. Pathways linking exposure and AOs can emerge from the complex metabolic web, i.e., the AOP reconstitution. Apart from the omic-wide associated studies, a 'meet-in-the-middle' strategy has been applied in univariate mediation analysis (Chadeau-Hyam et al., 2011; Vineis et al., 2013). MIMA can potentially uncover many hidden molecular events and link exposure to disease or outcome via the metabolic network (Figure 1) (Huang et al.,

2018; Huang et al., 2019; Liu et al., 2020b; Wang et al., 2019b; Wu et al., 2018), where the pathways embedded in the complex metabolomics can be identified. Given arsenic exposure is the risk of male infertility on the traditional "black box" epidemiology (Wang et al., 2016), one of the more recent application of the O2PLS model proposed in Figure 1 has tried to reconstitute the linkages between exposure biology of arsenic and pathology of male infertility via non-targeted analysis of urinary metabolome (Wu et al., 2021). In this case-control study, twelve metabolites each have been defined as the arsenic exposure biomarkers and infertile pathology biomarkers, respectively. Seven of them can directly bridge the gap between arsenic and infertility. Interestingly, the core metabolism correlation network further highlighted that testosterone is the vital hub to transfer the arsenic effects to infertile risk. The network indicated arsenic methylation that coupled disruption of one-carbon metabolism and oxidation stress and the adverse effects extended to fatty acid oxidation and steroidogenesis dysfunction.

MIMA research can be fundamentally supported by utilizing traditional epidemiological study designs. However, three aspects are critical to improving these studies:

1. Holistic analyses of molecular exposomes and endogenous metabolites.

2. Reconstructing the global endogenous metabolism network and recognizing the metabolism modules and hub nodes that respond to environmental factor exposure and reflect the AO, respectively. Additionally, profiling the pathways and assessing the biomarkers (with sound sensitivity and specificity).

3. Conduct causative inferences for biomarkers using mediation or moderation analysis. The identified biomarkers then return from holism to reductionism because they are deduced from a global view with fully mechanical information.

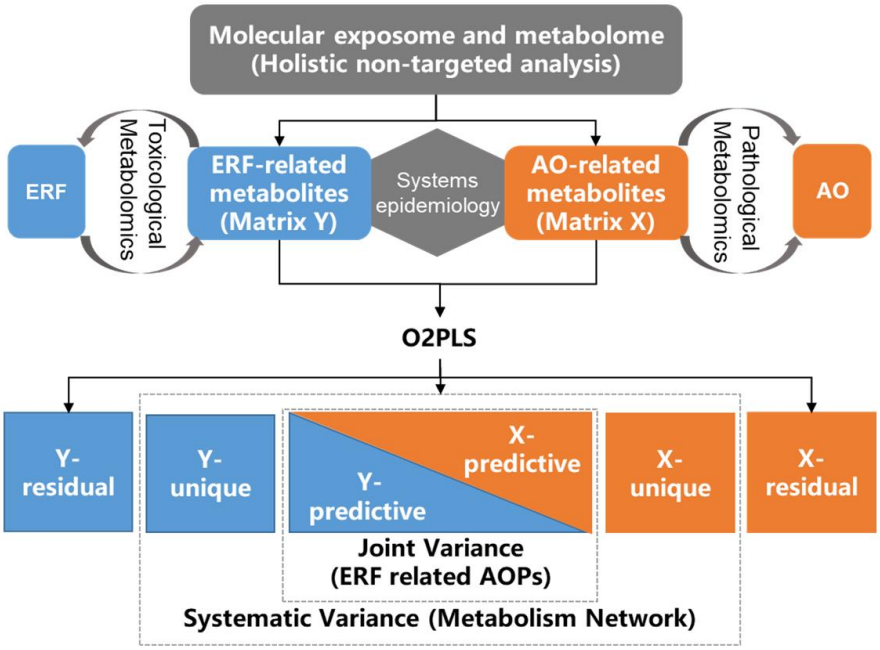

Figure 1: Meet-in-metabolite analysis model: O2PLS application to bridge the gap between environmental risk factors and adverse outcome.

Note: Biomonitoring and metabolomics can be integrated into top-down measurements under the MIMA framework and non-targeted analysis in holism is technical foundation of systems epidemiology. Quantitatively and qualitatively annotating endogenous and exogenous analytes and biologically reconstituting system dynamics concerning AOPs or networks are the major goals of AOP-oriented molecular epidemiology in MIMA. ERF = Environmental risk factor; AO = Adverse outcome; AOPs = Adverse Outcome Pathways; O2PLS = Two-way Orthogonal Partial Least-Square; Joint variance are the intersection of metabolite sets that represent both ERF-related and AO-related metabolites, i.e., Y-predictive and X-predictive, respectively. While Y-unique represent the metabolites are independent to AO and X-unique represent the metabolites are independent to ERF, respectively. Y-residual and X-residual are the unexplained parts of the total variations in the model.

Genomic-wide associated disease and human exposomics

Genome-wide association studies have revealed many genetic associations and mapped certain networks to improve our understanding of the nature of disease; however, these mappings only account for a small fraction of the disease risks and most parts can be the environment-wide associated (Rappaport and Smith, 2010). Utilizing exposomics to assess all individual exposures in a lifetime and how they relate to disease idealizes exposure measurements in epidemiology (Wild, 2005), emphasizing both systematic and accurate analysis characteristics. The National Institute of Environmental Health Sciences has defined the set of environmental exposures that shifts the human body condition from healthy by chemical exposures, diet, physical activity, stress, pre-existing disease, and the use of addictive substances. Considering human exposure biomonitoring, ethically available biological samples such as blood and urine from the investigated population contain a wide variety of global metabolome information including that of exposures. Based on metabolism, the related exposure responses and health outcomes or disease in cohorts can also be compared (Rappaport et al., 2014; Shen et al., 2014). Mathematically, human body is a cohesive conglomeration of interdependent components that are delineated via both spatial and temporal boundaries. Therefore, metabolomics that encompasses complete internal exposure molecular information and associates a disease can be modeled via MIMA.

Profiling exposure-oriented disease or adverse outcome on metabolism

Given the limited number of pathways for responding to exposure to various pollutants, exposure biology can be applied to address toxicologically related metabolites; along with the pathological metabolite analysis (Compton et al., 2019), how environmental factors initiate the toxic hazards and transformed into AO can be highlighted, which is vital for environmental disease prevention and medicine. For example, regarding arsenic effects (Zhang et al., 2014c) and male infertility (Huang et al., 2019; Shen et al., 2013; Wang et al., 2016; Zhang et al., 2014a; Zhang et al., 2014b), the pathways of arsenic methylation coupling one-carbon metabolism disruption together with oxidation stress can propagate to the fatty acid oxidation and steroidogenesis dysfunction indicated by testosterone has been recently profiled (Wu et

al., 2021). Although the change of testosterone, a key molecular event for male fertility, is not so sensitive and specific to both arsenic-related and male infertility-related metabolites (Wu et al., 2021; Zhang et al., 2014c), MIMA procedure by O2PLS dose has centralized its role in-between arsenic exposure and male infertility. Due to the nonlinear dynamics of biological systems, it is no surprise that not all molecular events are sensitive or specific enough to respond to exposure and/or AOs, which may typically be omitted in the traditional omic-wide associated biomarker investigations.

Apart from arsenic, the exposure-oriented metabolism of phthalates, perfluorinated compounds, air pollution of $PM_{2.5}$ have been profiled on human monitoring data of urine, blood or meconium samples, in which male infertility, gestational diabetes mellitus (GDM), chronic obstructive pulmonary disease (COPD) or alteration of cardiorespiratory function have been linked via AO-oriented metabolism (Huang et al., 2018; Huang et al., 2019; Liu et al., 2020b; Peng et al., 2015; Wang et al., 2017; Wang et al., 2019b; Wu et al., 2018). Interestingly, the cardiorespiratory effects of a very low ozone exposure (8.7 ± 6.6 ppb) that associated with changes in metabolic profiles among the vulnerable children can be observed, in which ozone below the current indoor standards was associated with the deteriorated cardiovascular function by disturbing bile acid and endogenous nitric oxide-related oxidation and inflammation, and associated with the exacerbated airway inflammation by reducing GPx-related anti-oxidation (Liu et al., 2021). The result implied that MIMA is comprehensive and powerful for tracing the molecular changes from exposure to outcome.

Although metabolome has encompassed all biological aspects of exposure, toxicology and pathology, the analytical aspects are still challenged by the holistic measurement of exogenous and/or endogenous metabolome, it is believed that totality but solitary of the risk factors are decisive in environment related disease pathology. Metabolomics usually can be conducted through high-resolution mass spectrometry (HRMS) with high throughput by non-targeted strategy. However, molecular events that are important for reconstituting the entire pathway from exposure to health outcome may be missed due to the concentration biased metabolome detection. Therefore a systematic analysis without bias towards any metabolite and/or pollutant is required. Additional flexible strategies must be adopted when all small molecules are intended to be monitored in viewpoint of metabolomics. Especially some analytical aspects in exposomics are challenging (Dennis et al., 2017), such as the missing of low level xenobiotics in non-targeted analysis, the identification of unknown small molecules without database, and the metabolomic heterogeneity in different human samples.

Systematic analysis of human exposure to environmental chemicals

We routinely face a man-made chemical world. Over 145,000 chemicals are registered by the European Union Registration, Evaluation, Authorization and Restriction of Chemical Substances regulation (https://echa.europa.eu/web/guest/information-on-chemicals/pre-registered-substances). A recently released update of the Toxic Substances Control Act (TSCA) inventory list for chemicals that are actively being manufactured by the U.S.

Environmental Protection Agency (EPA) as of February 19, 2019, demonstrated that less than half of the total number of chemicals on the current TSCA inventory list, i.e., 40,655 of the 86,228 (47%) chemicals, are currently being sold. Except for a few persistent organic chemicals (POPs), most of the high production volume (Nikfar et al., 2014) chemicals can be considered pseudo-persistent chemicals because of their continuous emission to the environment, even if their half-lives are short (Bergman et al., 2012; Daughton, 2003). These chemicals plus POPs can be a great source of concern if vulnerable subpopulations are exposed to amounts that result in adverse effects (Nicholson et al., 2004). The National Health and Nutrition Examination Survey, a well-known targeted biomonitoring study, includes a few hundred preselected chemicals for the survey. Therefore, it is an enormous gap between the ongoing biomonitoring and the current human exposure to environmental chemicals, which has hindered the exposure risk assessment for evaluating environmental health risks.

The EPA Non-Targeted Analysis Collaborative Trial (ENTACT) has used suspected screening for approximately 1,200 chemical substances from the EPA ToxCast library in house dust, silicone wristbands, and human serum (Ulrich et al., 2019) to systematically assess human exposure to environmental chemical pollutants. Approximately 40,153 to 114,100 small molecules are populated in human body (Wishart et al., 2018), with only 5,835 reconstructed by the endogenous human metabolism (Angione, 2019). It has been estimated that the four major small-molecule categories, namely endogenous chemicals, food chemicals, pollutants, and pharmaceuticals in blood, exhibit concentrations that cover a 10^7-fold range. One of the greatest challenges to screening and identifying unknown pollutants in human samples is their minuscule concentrations, which are typically thousand times lower (median of $2.4x10^{-4}$ μM) than those of endogenous chemicals (0.94 μM), food chemicals (1.00 μM), and pharmaceuticals (0.30 μM). Given that untargeted HRMS can detect >30,000 small molecule features in human serum (Ivanisevic et al., 2013), a platform bias results in measurement uncertainties for small molecules less than approximately 0.1 μM in 50-L of serum, where approximately 90% of pollutants and 30% of endogenous and food chemicals have been missed regarding the data-dependent acquisition of fragment ions (MS2). This is because only the top 10 most intense precursor ions (MS1) are subjected to MS2 (Wang et al., 2019a; Yan and Yan, 2015). When an expected system biomonitoring model (such as that of a top-down exposome (Rappaport, 2011) is applied, more flexible strategies should be applied to overcome the challenges in the nowadays non-targeted holistic analysis (Figure 2), such as data-dependent acquisition (DDA) approach.

Because of there is a huge concentration differences among the much lower pollutants and the relatively high endogenous chemicals, food chemicals, and drugs, respectively (Ivanisevic et al., 2013; Rappaport, 2014), results in many challenges to apply the holistic measurement to molecular exposome. DDA approaches can only select a few fractions of the most-intense (such as the top 10) mass spectra (MS) of precursor ions (MS1) subject to fragmentation (MS2) (Yan and Yan, 2015), which is the substantial bias for searching less abundant but biological important molecules in metabolomics. In addition to the commonly used DDA, the data-independent

acquisition (DIA) can theoretically address all molecules with MS/MS features, however, MS1 and MS2 ions might co-elute and difficult to identify the parent-daughter linkages for small molecules when compared to for proteins. As a result DIA of MS2 covers a broad range of precursor ion fragments and the link between MS1 and MS2 might get often lost. With the recent development of DIA protocols such as scanning sequential window acquisition of all theoretical spectra (SWATH) (Raetz et al., 2020), a novel DIA method SONAR (Juvvadi et al., 2018), and ion mobility MS/MS (IM-MS/MS) (Zheng et al., 2017; Zhou et al., 2018), MS1 selectivity have been greatly increased. The extra dimension for mass separation in the new DIA protocols of scanning SWATH, SONAR and IM–MS/MS can help to reconstitute MS1/MS2 parent-daughter linkages; then the enhanced unbiased measurement for particularly low abundance metabolites in non-targeted analysis can be readily obtained (Figure 2).

The challenge still remains to identify non-targeted small molecules. With only MS/MS features, many pollutants cannot be identified due to the database searching limitations. Currently, over 114,100 metabolites have been listed in the Human Metabolome Database (version 4.0) (Angione, 2019), but few are metabolites from xenobiotics. For example, meconium metabolites of methylepicatechin, methylxanthine, dimethyluric acid and vanilloylglycine are exogenous and are commonly present in green teas, red wine, cocoa products, and many fruits, which have been associated with GDM risk (Peng et al., 2015). Computational MS (Feunang et al., 2016) for small molecules should be further developed in combination with their biological transformations, such as the in-molecule glucuronide feature (Tsugawa et al., 2019; Walmsley et al., 2019). These kinds of features are commonly occurring for the metabolites of xenobiotics for assisting their excretion via urine and/or bile pathways but then computational-based exogenous metabolite molecular structure reconstruction (Scheubert et al., 2013) can become rational and will be applied to molecular exposome analysis. Considering sample preparation methods, sub-grouping strategies can be used to improve the non-targeted analysis of human exposome in the model of metabolome because after extraction, enrichment and purification different xenobiotics can be classified and collected by their properties such as lipophilicity, hydrophilicity or formation of adducts to other biomolecules (Figure 2).

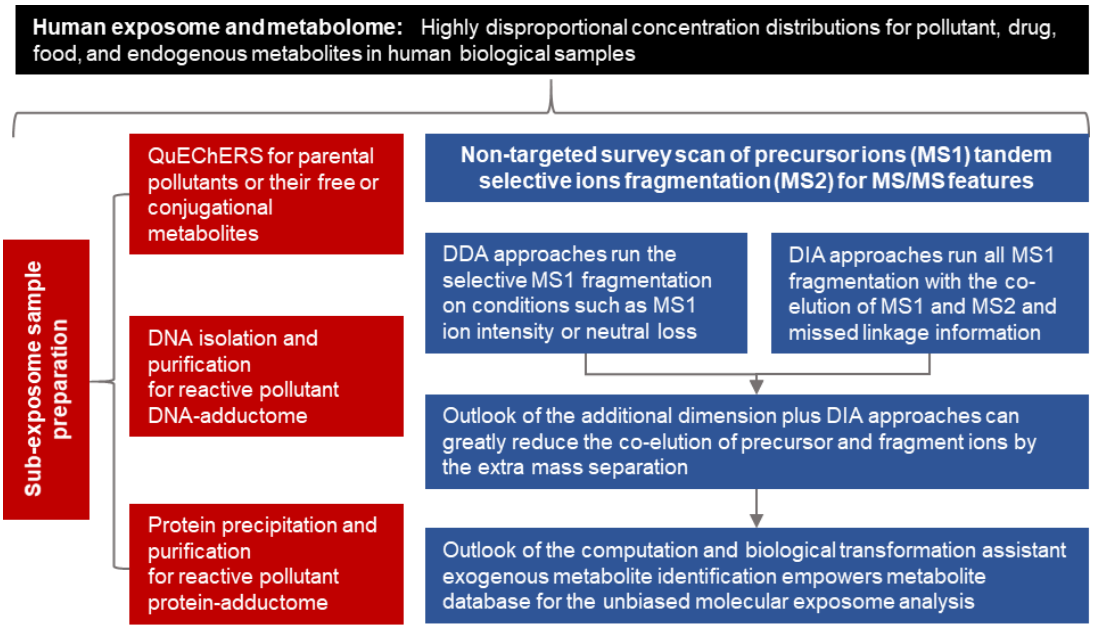

Figure 2. Suggested perspectives for a holistic non-targeted analysis of human molecular exposome.

Note: Various levels of small molecules of xenobiotics can be hidden in human metabolome and requires an integral strategy from initial sample preparation to final analytical method selection, which improves systems thinking of the aspects of analytical challenges in molecular exposomics.

Heterogeneity of environmental chemicals and exposure biomarkers in humans

Monitoring human exposure to suspected environmental chemicals using internal biomarkers has been accepted as the most accurate exposure assessment approach (Barr et al., 2005; Smolders et al., 2009). Internal chemicals and/or their metabolites should constitute the most abundant fractions and predominately represent the external exposure to be considered suitable biomarkers, where the measured analytes exhibit the sound specificity and sensitivity is also required to respond to their exposure. After entering the human body, xenobiotics are metabolized via the pronounced phase-I and phase-II reactions in the liver and kidney (or by gut microbiota when exposure via diet). In addition, certain exogenous chemicals can be transformed into the activated derivatives of protein and DNA adducts, and some adducts may initiate the immune system response, in which the xenobiotic-protein adducts can form antigen-antibody complexes (Pallardy and Bechara, 2017). Chemical pollutants have been categorized as POPs, non-persistent organic chemicals, bioaccumulative metals, non-bioaccumulative metals, and others (Smolders et al., 2009). Analyte can either surrogate the chronic exposure when a chemical is persistent or accumulated, or they can represent the acute exposure when it is reactive or non-bioaccumulative. Reactive chemical metabolic fractions excreted via urine may indicate acute exposure, or they may be classified as chronic exposure biomarkers when combined with proteins. Although few cases have been compared for environmental chemicals, the use of fasting plasma glucose (free blood glucose level) as acute biomarker and the hemoglobin A1c adduct (HbA1c) as chronic one have been well-documented (Association, 2017). Finally, the relative

abundance and time-dependent variation for the reactive chemicals as antigen-antibody complexes have not yet been addressed for biomonitoring.

Exposure biomarker selection is also subject to monitoring and molecular exposure life stages (Dennis et al., 2017). This is because the toxicokinectics (Liu et al., 2020a) and ethical convenience of sampling are development-specific. Various human sample types have been used as biomonitoring matrices (Smolders et al., 2009; Pallardy and Bechara, 2017). Placenta, cord blood, and meconium (Huang et al., 2019) samples are utilized to assess perinatal exposure, which may cover all chemical types. For example, meconium is formed by the fetus as early as the 12^{th} week of gestation and accumulates until birth; it is a repository of endogenous and exogenous agonists and metabolites, is capable of capturing disease-relevant metabolic profile changes and identifying novel biomarkers (Peng et al., 2015). However, the use of meconium must be further explored because the extraction for many organic pollutants in this matrix exhibits more difficulties than in placenta and cord blood. Sampling from newborns and infants is more difficult than sample collection from other developmental stages; therefore, the diaper urine and blood spots must be deeply mined for biomonitoring. For the two most convenient types of samples, blood and urine, persistent or bioaccumulative chemicals are likely found in the blood, while the reactive and non-bioaccumulative ones are likely to be concentrated in urine in their free forms as parental or their metabolites.

Subject to the chemical type, POP biomonitoring has been well-documented for targeted analytical strategies; the related biomarkers are typically the parental chemicals in serum samples (Smolders et al., 2009; Pallardy and Bechara, 2017). Certain persistent organohalogen compounds (OHCs) may be natural chemicals (Agarwal et al., 2014; Teuten et al., 2005). However, most of them are man-made, and a systematic analysis of these trace levels of POPs in humans can be conducted using both target and non-target approaches, such as gas chromatography (GC)-HRMS for semi-volatile and lipophilic POPs (Alonso et al., 2017; Goto et al., 2020; Yang et al., 2019) and liquid chromatography-HRMS for hydrophilic POPs such as perfluoroalkyl and polyfluoroalkyl substances (PFASs) (Concha-Grana et al., 2018; Ruan and Jiang, 2017).

The exposure biomarker selection for reactive pollutants is more challenging to acquire than that for persistent ones. Reactive chemicals and their metabolites may occur in free forms in metabolites through derivatization via functional groups such as -OH, -SH, -NH2, and -COOH, and/or by conjugation with glucuronic acid, sulfate, glutathione, or acetyl. Free metabolites and small molecular conjugates are typically concentrated in urine and are ready for excretion. Therefore, they are typically selected as biomarkers to assess acute exposure to reactive pollutants. The electrophilic activated chemicals can also covalently adduct to nucleophilic sites in proteins (such as the sulfhydryl group of cysteine, ε-amino group of lysine, and imidazole group of histidine in proteins (Shibata and Uchida, 2019) in hemoglobin and albumin) and DNA (Cooke et al., 2018; Guo and Turesky, 2019). When measuring these derivatives in adductomics (Rappaport et al., 2012), hemoglobin adducts and human serum albumin (HSA) are preferable to those of DNA and glutathione for characterizing chronic exposure because of their greater abundance and longer residence times in human blood (Needham and

Sexton, 2000). Consequently, these circulating protein adducts can be used as chronic exposure biomarkers for reactive chemicals. In addition to free protein adducts, circulating antigen-antibody complexes (Aguirre-Gamboa et al., 2016; Sharma et al., 2017) may pool the pro-/pre- or haptenic xenobiotics for immunity-oriented assessment.

Metal biomarkers are comparable to organic chemicals, including elements or their metabolic species in urine and blood samples. The common approach to address metal speciation in biosciences is called metallomics, and various metals and species can be sensitively measured by LC-inductively coupled plasma (ICP)-MS. Urinary and serum samples are typically digested and measured by ICP-MS or ICP-optical emission spectroscopy, where the elements can be applied as biomarkers to indicate their total exposure status (Medda et al., 2016; Troisi et al., 2019). Certain metal or metalloid species, such as methylated arsenic and mercury, can be measured in urine and blood samples by LC-ICP-MS. Many small (oxalate, citrate, tartrate, amino acids, and oligopeptides) and large (nucleic acids, polysaccharides, and proteins) biomolecules can bind to metals, generating various chemical species (Lopez-Barea and Gomez-Ariza, 2006). Similar to organic chemicals, metals can occur in free forms, small metabolic species, or metalloproteomic forms with proteins (Coverdale et al., 2019). For example, serum albumin is a highly abundant plasma protein associated with the transport of metal ions (Curry et al., 1998); therefore, HSA can be used in the exposome proteomic forms to metals (Curry et al., 1998; Varshney et al., 2010). The inducible metallothioneins are Cys-rich metal-binding proteins, and they can be used as both exposure and stress biomarkers for metal exposomics (Coverdale et al., 2019).

Screening strategies for parental chemicals and their free form metabolites

Monitoring pre-selected target chemical pollutants (including their demarcated metabolites) in various human samples and their applications throughout the life stages has been sufficiently performed (Smolders et al., 2009). In clinical and forensic toxicological fields, the parental chemicals or their defined metabolites can be screened in the pre-selected list as suspected unknowns in human samples. The typical strategy is the chemical parental structure-dependent screening. For the suspected unknown OHC screening, sample preparation procedures and GC-MS analysis have determined chemicals of neutral, lipophilic, and semi-volatile properties, which exhibit well-demarcated chemical subclasses and can match a holistic screening. With the assistance of sample clean-up and target concentration, non-target OHC screening was performed using a full-scan screening on the GC×GC-HRToFMS (resolution power >5000 with mass error ±10 ppm) and qualitative analysis by GC-magnetic-sector HRMS (resolution power 8000-12000 with mass error <10 ppm) (Goto et al., 2020), and over 300 OHCs were identified in marine dolphins (Alonso et al., 2017). Over 4,000 PFASs have been suspected to enter the environment (Ritscher et al., 2018). PFASs contain at least one perfluoroalkyl moiety (C_nF_{2n}), this common feature is useful for screening their human exposure. The systematic measurement of PFASs (Shibata and Uchida, 2019) is urgent and can be acquired using LC-time-of-flight or Orbitrap HRMS, which is equipped with electrospray ionization for ionic PFASs and atmospheric pressure photoionization for neutral PFASs. The proportions of unidentified organofluorines

rose from approximately 20% to 50% in German plasma samples collected during 2000-2009, indicating human exposure to various unknown PFASs (Yeung and Mabury, 2016). Holistic screening can be performed on the perfluoroalkyl moiety -C_nF_{2n}- through a combination of criteria such as mass balance (Shibata and Uchida, 2019). Regarding the in-molecule diagnostic features of [C2F5]- (m/z 118.992) and [C_3F_7]- (m/z 168.988), low levels of unknown C5-C17 poly- and perfluoroalkyl substances have been identified in water (Liu et al., 2015). The data requirements include both DDA and DIA modes in PFAS HRMS target and non-target analyses. These strategies can also be used for other chemicals.

The pre-selected suspected unknown screening for free reactive chemicals and their metabolites can be made available when the molecular structures of the biomarkers in blood or urine have been documented, in which the unknown screening procedures (Plassmann et al., 2015) can be referenced with additional sample preparation procedures. For example, the quick, easy, cheap, rugged, and safe preparation method can be used to extract compounds covering a broad domain for GC or LC–MS analysis (Perestrelo et al., 2019). To extend the suspected unknown list, the ENTACT has attempted to advance xenobiotic chemical analysis in environmental and biological media by using the ToxCast library of chemical substances, DSSTox database, and CompTox Chemicals Dashboard (Ulrich et al., 2019). However, free reactive chemicals and their metabolites may only account for the small fractions of these pollutants, and most of them should have been biologically transformed. Thus, the pollutant dependent screening may only be suitable for the small subclasses of POPs.

Screening strategies for biologically transformed environmental chemicals

For reactive chemical pollutants, the metabolism biological basis involves converting lipid-soluble, non-polar, and non-excretive forms to water-soluble and polar forms that are excretive in bile and urine. Their metabolites may exist in free forms (by exposing or adding functional groups) or in conjugated forms (to large water-soluble and charged endogenous molecules), all of which are ready for excretion via urine, and their urinary concentration can indicate acute exposure. Certain fractions may also form adducts with DNA and proteins and exist in the blood circulation; their half-lives communicate the related macromolecule degradation and can therefore indicate certain chronic exposure types. The identification of a reactive chemical that adducts/conjugates to an endogenous biomolecule can be assisted by using in-molecule diagnostic features (i.e., the endogenous sub-structures) (Plassmann et al., 2015).

Urinary sub-exposome of conjugates: Xenobiotic reactions primarily occur in the liver (hepatocytes) and occur less frequently in the kidney (proximal tubular cells), lungs (Clara cells and alveolar cells), intestines (mucosa lining cells), skin (epithelial cells), and testes (seminiferous tubules and Sertolis cells). Molecular sub-exposome of urinary metabolites of the four major phase II metabolism conjugations (glucuronic acid, sulfate, glutathione, and acetyl) can be separately profiled, where the sulfates, glucuronide, acetyl, and mercapturic acids (Frigerio et al., 2020; Yao et al., 2016) may be used as in-molecule flagging features for the systemic unknown screening, respectively. In addition, the human metabolome (Walker et al., 2019) and related

databases can be used to aid the screening process. Because of the mass spectral acquisition in parallel modes cycles back and forth, where the MS1 of the full scan with ultrahigh resolving power (such as RP = 120,000, mass accuracy ≤3 ppm) and the MS2 with in-source fragmentation flagging scans to flag the suspected target retention times using diagnostic fragments, the lower xenobiotic concentrations can be theoretically addressed. However, challenges remain for computation-based chemical identification (Bocker and Duhrkop, 2016; Dührkop et al., 2013; Hufsky et al., 2014; Ruttkies et al., 2016; Scheubert et al., 2013).

Sub-exposome of adducts: When the reacted xenobiotic species are trapped by macromolecules in the targeting tissue or circulation system, reactions can occur at the genome or proteome scale. To improve the analytical efficiency, protein and DNA sample preparation and purification can be applied before their digestion. Similar platforms in metabolome analysis can then be used for adducts. LC−MS^n-based DNA adductomic investigations can utilize a common structural feature of deoxyribonucleosides, in which a deoxyribose moiety bound to the nucleobase through a glycosidic bond (Balbo et al., 2014). The DIA wide selective ion monitoring/MS2 methodology (Guo et al., 2017) with HRMS can detect many DNA adducts through non-targeted screening and computational data analysis. Reactive organic pollutants most often bind covalently through their electrophilic properties to react with proteins when they are trapped by circulating proteins (Li et al., 2011b), in which the nucleophilic protein sites added to electrophilic toxicants. Many reactions can be observed in hemoglobin and HSA or toxic-targeting DNA (Kanaly et al., 2006). An adductomic strategy can then be directly applied for the suspected xenobiotic biomonitoring. For proteins and nucleic acids that are biologically degraded, the embedding xenobiotics may be measured as metabolites along with metabolome scale reactive chemicals (Li et al., 2011a; Ma and Subramanian, 2006).

Circulating antigene-antibody complexes may serve as another sub-exposome protein pool for reactive chemicals. The innate and adaptive immune systems have been developed by living organisms to protect them from "outside" viruses, bacteria, and parasites. Thus, "outside" chemicals can be translated into innate immune system activation, which may be the consequence of different key steps that allow dendritic cells (DCs) to initiate immune system adaptation (Association, 2017). The human immune system is highly reactive to the environment, and 80% of the measured immunological parameters are affected by the environment (≥50% variance) (Brodin et al., 2015). Many environmental chemicals, acting as haptens, can bind to HSA and cause the immune system to misidentify self-tissue as an invader and launch an immune response against it (autoimmunity) (Vojdani et al., 2015). Apart from the liver and other metabolic organs, innate immune cells such as monocytes, macrophages, dendritic cells, and polymorphonuclear cells play a fundamental role in xenobiotic metabolism. Xenobiotic metabolic conversion was observed in dermal Langerhans cells (immature DCs), containing cytochrome P4501A (CYP1A) enzymes. After conversion by the CYP1A enzyme, xenobiotics form a complex with self-proteins, which are processed and presented as major histocompatibility complex class I and class II molecules. The xenobiotic–protein complex presented by antigen-presenting cells is subsequently

recognized by T cells, which help B cells with antibody production. The conjugation of xenobiotics to self-proteins makes them highly immunogenic and therefore elicits the production of anti-xenobiotic antibodies, which play a physiological role in clearing xenobiotics from the body (Association, 2017; Sajid and Agrewala, 2019).

Many heavy metal pollutants such as mercury, nickel, and cobalt can react with oxidized proteins to form protein metal chelate complexes (Sajid and Agrewala, 2019). These xenobiotics may be indirectly measured via antigen-antibody complexes, which are multi-molecular complexes that are typically stabilized via the reversible interaction of static electricity, hydrogen bonds, or the van der Waals force. These small pollutants directly bind to self-proteins or bind indirectly after hepatic or extrahepatic conversion from prohaptens to haptens, generating hapten-protein adducts (Sajid and Agrewala, 2019). When comparing free HSA, circulating antigen-antibody complexes (Sharma et al., 2017; Medda et al., 2016), such as the aflatoxin-HSA adduct, which can elevate IgG and/or IgM (Vojdani et al., 2003), may be concentrated, and xenobiotic pooling should be investigated through immunoprecipitation using HRMS. For example, MS methods revealed that an extreme sensitizer 2,4-dinitro-1-chlorobenzene, and methylchloroisothiazolinone modified a greater number of nucleophilic HSA sites than the moderate sensitizer cinnamaldehyde. However, the weak/non-sensitizer 6-methyl coumarin was restricted to a single nucleophilic residue within HAS (Esser and Martin, 2017).

In summary, the metabolome encompasses all exposure burden information and exhibits greater potential to generate data to enhance exposure assessment regarding exposomics than that of any other omics research. In addition, lifetime dimensional information can be obtained by repeatedly measuring ethically available life-staged blood and urine. From the human bio-monitoring perspective, these biological samples have convened all possible global postnatal information on the investigated population, including chemical exposure, biological response, and the potential linkage to a defined health outcome. Regarding maternal blood and urine during pregnancy, partly attributed ‘prenatal’ exposure cord blood, placenta, and meconium may ethically be available to characterize fetal prenatal conditions. Finally, the metabolome is useful for conducting biology-based estimates of individual and public health risks.

Acknowledgements: This review have been developed and structured to resume our previous works from 2014 or earlier. We sincerely thank the National Key R&D Program of China (2017YFC0211602， 2017YFC0211600) and National Natural Science Foundation of China (No. 21177123, 91543113, 21777157, 22076158) for the support. We also would like to thank Editage (www.editage.com) for English language editing.

Agarwal, V., El Gamal, A.A., Yamanaka, K., Poth, D., Kersten, R.D., Schorn, M., Allen, E.E.,

Moore, B.S., 2014. Biosynthesis of polybrominated aromatic organic compounds by marine

bacteria. Nat. Chem. Biol. 10 (8), 640-U182.

Aguirre-Gamboa, R., Joosten, I., Urbano, P.C.M., van der Molen, R.G., van Rijssen, E., van Cranenbroek, B., Oosting, M., Smeekens, S., Jaeger, M., Zorro, M., Withoff, S., van Herwaarden, A.E., Sweep, F., Netea, R.T., Swertz, M.A., Franke, L., Xavier, R.J., Joosten, L.A.B., Netea, M.G., Wijmenga, C., Kumar, V., Li, Y., Koenen, H., 2016. Differential Effects of Environmental and Genetic Factors on T and B Cell Immune Traits. Cell Reports 17 (9), 2474-2487.

Alonso, M.B., Maruya, K.A., Dodder, N.G., Lailson-Brito, J., Azevedo, A., Santos-Neto, E., Torres, J.P.M., Malm, O., Hoh, E., 2017. Nontargeted Screening of Halogenated Organic Compounds in Bottlenose Dolphins (Tursiops truncatus) from Rio de Janeiro, Brazil. Environ. Sci. Technol. 51 (3), 1176-1185.

Angione, C.J.B.R.I., 2019. Human Systems Biology and Metabolic Modelling: A Review—From Disease Metabolism to Precision Medicine. 2019.

Association, A.D.J.D.C., 2017. 2. Classification and Diagnosis of Diabetes. 40 (Suppl 1), S11.

Balbo, S., Turesky, R.J., Villalta, P.W., 2014. DNA Adductomics. Chem. Res. Toxicol. 27 (3), 356-366.

Barr, D.B., Wang, R.Y., Needham, L.L., 2005. Biologic monitoring of exposure to environmental chemicals throughout the life stages: Requirements and issues for consideration for the National Children's Study. Environ Health Persp 113 (8), 1083-1091.

Bergman, A., Heindel, J., Jobling, S., Kidd, K., Zoeller, R.T., 2012. State-of-the-science of endocrine disrupting chemicals, 2012. Toxicol. Lett. 211, S3-S3.

Bocker, S., Duhrkop, K., 2016. Fragmentation trees reloaded. J. Cheminformatics 8, 26.

551 Brodin, P., Jojic, V., Gao, T.X., Bhattacharya, S., Angel, C.J.L., Furman, D., Shen-Orr, S.,
552 Dekker, C.L., Swan, G.E., Butte, A.J., Maecker, H.T., Davis, M.M., 2015. Variation in the Human
553 Immune System Is Largely Driven by Non-Heritable Influences. Cell 160 (1-2), 37-47.
554 Chadeau-Hyam, M., Athersuch, T.J., Keun, H.C., De Iorio, M., Ebbels, T.M.D., Jenab, M.,
555 Sacerdote, C., Bruce, S.J., Holmes, E., Vineis, P., 2011. Meeting-in-the-middle using metabolic
556 profiling - a strategy for the identification of intermediate biomarkers in cohort studies.
557 Biomarkers 16 (1), 83-88.
558 Compton, C.C., Robb, J.A., Anderson, M.W., Berry, A.B., Birdsong, G.G., Bloom, K.J., Branton,
559 P.A., Crothers, J.W., Cushman-Vokoun, A.M., Hicks, D.G., Khoury, J.D., Laser, J., Marshall,
560 C.B., Misialek, M.J., Natale, K.E., Nowak, J.A., Olson, D., Pfeifer, J.D., Schade, A., Vance, G.H.,
561 Walk, E.E., Yohe, S.L., 2019. Preanalytics and Precision Pathology: Pathology Practices to
562 Ensure Molecular Integrity of Cancer Patient Biospecimens for Precision Medicine. Arch Pathol
563 Lab Med 143 (11), 1346-1363.
564 Concha-Grana, E., Fernandez-Martinez, G., Lopez-Mahia, P., Prada-Rodriguez, D.,
565 Muniategui-Lorenzo, S., 2018. Fast and sensitive determination of per- and polyfluoroalkyl
566 substances in seawater. J. Chromatogr. A 1555, 62-73.
567 Cooke, M.S., Hu, C.W., Chang, Y.J., Chao, M.R., 2018. Urinary DNA adductomics - A novel
568 approach for exposomics. Environ. Int. 121, 1033-1038.
569 Coverdale, J.P.C., Barnett, J.P., Adamu, A.H., Griffiths, E.J., Stewart, A.J., Blindauer, C.A.,
570 2019. A metalloproteomic analysis of interactions between plasma proteins and zinc: elevated
571 fatty acid levels affect zinc distribution. Metallomics 11 (11), 1805-1819.
572 Curry, S., Mandelkow, H., Brick, P., Franks, N., 1998. Crystal structure of human serum albumin

573 complexed with fatty acid reveals an asymmetric distribution of binding sites. Nat. Struct. Biol.
574 5 (9), 827-835.
575 Daughton, C.G., 2003. Cradle-to-cradle stewardship of drugs for minimizing their environmental
576 disposition while promoting human health. II. Drug disposal, waste reduction, and future
577 directions. Environ. Health Perspect. 111 (5), 775-785.
578 Dennis, K.K., Marder, E., Balshaw, D.M., Cui, Y.X., Lynes, M.A., Patti, G.J., Rappaport, S.M.,
579 Shaughnessy, D.T., Vrijheid, M., Barr, D.B., 2017. Biomonitoring in the Era of the Exposome.
580 Environ. Health Perspect. 125 (4), 502-510.
581 Dührkop, K., Scheubert, K., Böcker, S.J.M., 2013. Molecular formula identification with SIRIUS.
582 3 (2), 506-516.
583 Esser, P.R., Martin, S.F., 2017. Pathomechanisms of Contact Sensitization. Curr. Allergy
584 Asthma Rep. 17 (12), 10.
585 Frigerio, G., Mercadante, R., Campo, L., Polledri, E., Boniardi, L., Olgiati, L., Missineo, P., Nash,
586 W.J., Dunn, W.B., Fustinoni, S., 2020. Urinary biomonitoring of subjects with different smoking
587 habits. Part II: an untargeted metabolomic approach and the comparison with the targeted
588 measurement of mercapturic acids. Toxicol. Lett. 329, 56-66.
589 Goto, A., Tue, N.M., Isobe, T., Takahashi, S., Tanabe, S., Kunisue, T., 2020. Nontarget and
590 Target Screening of Organohalogen Compounds in Mussels and Sediment from Hiroshima Bay,
591 Japan: Occurrence of Novel Bioaccumulative Substances. Environ. Sci. Technol. 54 (9), 5480-
592 5488.
593 Guo, J., Turesky, R.J.J.H.-t., 2019. Emerging technologies in mass spectrometry-based DNA
594 adductomics. 8 (2), 13.

595 Guo, J., Villalta, P.W., Turesky, R.J., 2017. Data-Independent Mass Spectrometry Approach
596 for Screening and. Identification of DNA Adducts. Anal. Chem. 89 (21), 11728-11736.
597 Helzlsouer, K.J., 1993. MOLECULAR EPIDEMIOLOGY - PRINCIPLES AND PRACTICES -
598 SCHULTE,PA, PERERA,FP. Science 262 (5136), 1082-1083.
599 Huang, Q.Y., Hu, D.Y., Wang, X.F., Chen, Y.H., Wu, Y., Pan, L., Li, H.Y., Zhang, J., Deng, F.R.,
600 Guo, X.B., Shen, H.Q., 2018. The modification of indoor PM2.5 exposure to chronic obstructive
601 pulmonary disease in Chinese elderly people: A meet-in-metabolite analysis. Environ. Int. 121,
602 1243-1252.
603 Huang, Q.Y., Liu, L.P., Wu, Y., Wang, X.F., Luo, L.Z., Nan, B.R., Zhang, J., Tian, M.P., Shen,
604 H.Q., 2019. Seminal plasma metabolites mediate the associations of multiple environmental
605 pollutants with semen quality in Chinese men. Environ. Int. 132, 11.
606 Hufsky, F., Scheubert, K., Bocker, S., 2014. Computational mass spectrometry for small-
607 molecule fragmentation. Trac-Trends Anal. Chem. 53, 41-48.
608 Ivanisevic, J., Zhu, Z.J., Plate, L., Tautenhahn, R., Chen, S., O'Brien, P.J., Johnson, C.H.,
609 Marletta, M.A., Patti, G.J., Siuzdak, G., 2013. Toward 'Omic Scale Metabolite Profiling: A Dual
610 Separation-Mass Spectrometry Approach for Coverage of Lipid and Central Carbon
611 Metabolism. Anal. Chem. 85 (14), 6876-6884.
612 Juvvadi, P.R., Moseley, M.A., Hughes, C.J., Soderblom, E.J., Lennon, S., Perkins, S.R.,
613 Thompson, J.W., Geromanos, S.J., Wildgoose, J., Richardson, K., Langridge, J.I., Vissers,
614 J.P.C., Steinbach, W.J., 2018. Scanning Quadrupole Data-Independent Acquisition, Part B:
615 Application to the Analysis of the Calcineurin-Interacting Proteins during Treatment of
616 Aspergillus fumigatus with Azole and Echinocandin Antifungal Drugs. J. Proteome Res. 17 (2),

780-793.

Kanaly, R.A., Hanaoka, T., Sugimura, H., Toda, H., Matsui, S., Matsuda, T., 2006. Development of the adductome approach to detect DNA damage in humans. Antioxid. Redox Signal. 8 (5-6), 993-1001.

Li, F., Lu, J., Ma, X.C., 2011a. Profiling the Reactive Metabolites of Xenobiotics Using Metabolomic Technologies. Chem. Res. Toxicol. 24 (5), 744-751.

Li, H., Grigoryan, H., Funk, W.E., Lu, S.S., Rose, S., Williams, E.R., Rappaport, S.M., 2011b. Profiling Cys(34) Adducts of Human Serum Albumin by Fixed-Step Selected Reaction Monitoring. Mol. Cell. Proteomics 10 (3), 13.

Liu, L.P., Wang, H., Li, X.Y., Tian, M.P., Huang, Q.Y., Zhang, J., Pan, H., Wen, K., Huang, Q.S., Yan, J.B., Tong, Z.D., Zhang, Y.L., Zhang, T.J., Zhang, Y.Y., Li, B., Wang, T., Shen, H.Q., 2020a. Infantile phthalate metabolism and toxico/pharmacokinetic implications within the first year of life. Environ. Int. 144, 12.

Liu, S., Huang, Q., Zhang, X., Dong, W., Zhang, W., Wu, S., Yang, D., Nan, B., Zhang, J., Shen, H., Guo, X., Deng, F., 2021. Cardiorespiratory Effects of Indoor Ozone Exposure Associated with Changes in Metabolic Profiles among Children: A Repeated-Measure Panel Study. The Innovation 2 (1), 100087.

Liu, S., Huang, Q.Y., Wu, Y., Song, Y., Dong, W., Chu, M.T., Yang, D., Zhang, X., Zhang, J., Chen, C., Zhao, B., Shen, H.Q., Guo, X.B., Deng, F.R., 2020b. Metabolic linkages between indoor negative air ions, particulate matter and cardiorespiratory function: A randomized, double-blind crossover study among children. Environ. Int. 138, 12.

Liu, Y.N., Pereira, A.D., Martin, J.W., 2015. Discovery of C-5-C-17 Poly- and Perfluoroalkyl

639 Substances in Water by In-Line SPE-HPLC-Orbitrap with In-Source Fragmentation Flagging.
640 Anal. Chem. 87 (8), 4260-4268.
641 Lopez-Barea, J., Gomez-Ariza, J.L., 2006. Environmental proteomics and metallomics.
642 Proteomics 6, S51-S62.
643 Ma, S.G., Subramanian, R., 2006. Detecting and characterizing reactive metabolites by liquid
644 chromatography/tandem mass spectrometry. J. Mass Spectrom. 41 (9), 1121-1139.
645 Medda, E., Minoprio, A., Nistico, L., Bocca, B., Simonelli, V., D'Errico, M., Calcagnile, A.,
646 Giuliani, A., Toccaceli, V., Minghetti, L., Alimonti, A., Stazi, M.A., Mazzei, F., Dogliotti, E., 2016.
647 The response to oxidative stress and metallomics analysis in a twin study: The role of the
648 environment. Free Radic. Biol. Med. 97, 236-243.
649 Needham, L.L., Sexton, K., 2000. Assessing children's exposure to hazardous environmental
650 chemicals: an overview of selected research challenges and complexities - Introduction and
651 overview. J. Expo. Anal. Environ. Epidemiol. 10 (6), 611-629.
652 Nicholson, A., Sandler, J., Seidle, T., 2004. An evaluation of the US High Production Volume
653 (HPV) chemical-testing programme: A study in (Ir)relevance, redundancy and retro thinking.
654 ATLA-Altern. Lab. Anim. 32, 335-341.
655 Nikfar, S., Behboudi, A., van der Kolk, J., Benson, A., 2014. High Production Volume (HPV)
656 Chemicals.
657 Pallardy, M., Bechara, R., 2017. Chemical or Drug Hypersensitivity: Is the Immune System
658 Clearing the Danger? Toxicol. Sci. 158 (1), 14-22.
659 Peng, S., Zhang, J., Liu, L., Zhang, X., Huang, Q., Alamdar, A., Tian, M., Shen, H., 2015.
660 Newborn meconium and urinary metabolome response to maternal gestational diabetes

mellitus: a preliminary case-control study. J Proteome Res 14 (4), 1799-1809.

Perestrelo, R., Silva, P., Porto-Figueira, P., Pereira, J.A.M., Silva, C., Medina, S., Camara, J.S., 2019. QuEChERS - Fundamentals, relevant improvements, applications and future trends. Anal. Chim. Acta 1070, 1-28.

Plassmann, M.M., Schmidt, M., Brack, W., Krauss, M., 2015. Detecting a wide range of environmental contaminants in human blood samples-combining QuEChERS with LC-MS and GC-MS methods. Anal. Bioanal. Chem. 407 (23), 7047-7054.

Raetz, M., Bonner, R., Hopfgartner, G., 2020. SWATH-MS for metabolomics and lipidomics: critical aspects of qualitative and quantitative analysis. Metabolomics 16 (6), 14.

Rappaport, S.M., 2011. Implications of the exposome for exposure science. J. Expo. Sci. Environ. Epidemiol. 21 (1), 5-9.

Rappaport, S.M., Barupal, D.K., Wishart, D., Vineis, P., Scalbert, A., 2014. The Blood Exposome and Its Role in Discovering Causes of Disease. Environ. Health Perspect. 122 (8), 769-774.

Rappaport, S.M., Li, H., Grigoryan, H., Funk, W.E., Williams, E.R., 2012. Adductomics: Characterizing exposures to reactive electrophiles. Toxicol. Lett. 213 (1), 83-90.

Rappaport, S.M., Smith, M.T., 2010. Environment and Disease Risks. Science 330 (6003), 460-461.

Ritscher, A., Wang, Z.Y., Scheringer, M., Boucher, J.M., Ahrens, L., Berger, U., Bintein, S., Bopp, S.K., Borg, D., Buser, A.M., Cousins, I., DeWitt, J., Fletcher, T., Green, C., Herzke, D., Higgins, C., Huang, J., Hung, H., Knepper, T., Lau, C.S., Leinala, E., Lindstrom, A.B., Liu, J.X., Miller, M., Ohno, K., Perkola, N., Shi, Y.L., Haug, L.S., Trier, X., Valsecchi, S., van der Jagt, K.,

683 Vierke, L., 2018. Zurich Statement on Future Actions on Per - and Polyfluoroalkyl Substances
684 (PFASs). Environ. Health Perspect. 126 (8), 5.
685 Ruan, T., Jiang, G.B., 2017. Analytical methodology for identification of novel per- and
686 polyfluoroalkyl substances in the environment. Trac-Trends Anal. Chem. 95, 122-131.
687 Ruttkies, C., Schymanski, E.L., Wolf, S., Hollender, J., Neumann, S., 2016. MetFrag relaunched:
688 incorporating strategies beyond in silico fragmentation. J. Cheminformatics 8, 16.
689 Sajid, M., Agrewala, J.N., 2019. Low prevalence of anti-xenobiotic antibodies among the
690 occupationally exposed individuals is associated with a high risk of cancer. Cancer Med. 8 (1),
691 246-260.
692 Scheubert, K., Hufsky, F., Bocker, S., 2013. Computational mass spectrometry for small
693 molecules. J. Cheminformatics 5, 24.
694 Sharma, M., Zhang, X., Zhang, S.M., Niu, L., Ho, S.M., Chen, A.M., Huang, S.X., 2017.
695 Inhibition of endocytic lipid antigen presentation by common lipophilic environmental pollutants.
696 Sci Rep 7, 13.
697 Shen, H., Xu, W., Zhang, J., Chen, M., Martin, F.L., Xia, Y., Liu, L., Dong, S., Zhu, Y.G., 2013.
698 Urinary metabolic biomarkers link oxidative stress indicators associated with general arsenic
699 exposure to male infertility in a han chinese population. Environ Sci Technol 47 (15), 8843-
700 8851.
701 Shen, H.Q., Xu, W.P., Peng, S.Y., Scherb, H., She, J.W., Voigt, K., Alamdar, A., Schramm,
702 K.W., 2014. Pooling samples for "top-down" molecular exposomics research: the methodology.
703 Environ. Health 13, 8.
704 Shibata, T., Uchida, K., 2019. Protein adductomics: A comprehensive analysis of protein

705 modifications by electrophiles. Free Radic. Biol. Med. 144, 218-222.
706 Smolders, R., Schramm, K.W., Stenius, U., Grellier, J., Kahn, A., Trnovec, T., Sram, R.,
707 Schoeters, G., 2009. A Review on the Practical Application of Human Biomonitoring in
708 Integrated Environmental Health Impact Assessment. J. Toxicol. Env. Health-Pt b-Crit. Rev. 12
709 (2), 107-123.
710 Teuten, E.L., Xu, L., Reddy, C.M., 2005. Two abundant bioaccumulated halogenated
711 compounds are natural products. (vol 307, pg 917, 2005). Science 308 (5727), 1413-1413.
712 Troisi, J., Giugliano, L., Sarno, L., Landolfi, A., Richards, S., Symes, S., Colucci, A., Maruotti,
713 G., Adair, D., Guida, M., Martinelli, P., Guida, M., 2019. Serum metallome in pregnant women
714 and the relationship with congenital malformations of the central nervous system: a case-control
715 study. BMC Pregnancy Childbirth 19 (1), 11.
716 Ulrich, E.M., Sobus, J.R., Grulke, C.M., Richard, A.M., Newton, S.R., Strynar, M.J., Mansouri,
717 K., Williams, A.J., 2019. EPA's non-targeted analysis collaborative trial (ENTACT): genesis,
718 design, and initial findings. Anal. Bioanal. Chem. 411 (4), 853-866.
719 Varshney, A., Sen, P., Ahmad, E., Rehan, M., Subbarao, N., Khan, R.H., 2010. Ligand Binding
720 Strategies of Human Serum Albumin: How Can the Cargo be Utilized? Chirality 22 (1), 77-87.
721 Vineis, P., van Veldhoven, K., Chadeau-Hyam, M., Athersuch, T.J., 2013. Advancing the
722 application of omics-based biomarkers in environmental epidemiology. Environ. Mol. Mutagen.
723 54 (7), 461-467.
724 Vojdani, A., Kharrazian, D., Mukherjee, P.S., 2015. Elevated levels of antibodies against
725 xenobiotics in a subgroup of healthy subjects. J. Appl. Toxicol. 35 (4), 383-397.
726 Vojdani, A., Pangborn, J.B., Vojdani, E., Cooper, E.L., 2003. Infections, toxic chemicals and

727 dietary peptides binding to lymphocyte receptors and tissue enzymes are major instigators of
728 autoimmunity in autism. Int. J. Immunopathol. Pharmacol. 16 (3), 189-199.
729 Walker, D.I., Valvi, D., Rothman, N., Lan, Q., Miller, G.W., Jones, D.P., 2019. The Metabolome:
730 a Key Measure for Exposome Research in Epidemiology. Curr. Epidemiol. Rep. 6 (2), 93-103.
731 Wang, R.H., Yin, Y.D., Zhu, Z.J., 2019a. Advancing untargeted metabolomics using data-
732 independent acquisition mass spectrometry technology. Anal. Bioanal. Chem. 411 (19), 4349-
733 4357.
734 Wang, X., Liu, L., Zhang, W., Zhang, J., Du, X., Huang, Q., Tian, M., Shen, H., 2017. Serum
735 metabolome biomarkers associate low-level environmental perfluorinated compound exposure
736 with oxidative /nitrosative stress in humans. Environ Pollut 229, 168-176.
737 Wang, X., Zhang, J., Xu, W., Huang, Q., Liu, L., Tian, M., Xia, Y., Zhang, W., Shen, H., 2016.
738 Low-level environmental arsenic exposure correlates with unexplained male infertility risk. Sci
739 Total Environ 571, 307-313.
740 Wang, Y.X., Wu, Y., Chen, H.G., Duan, P., Wang, L., Shen, H.Q., Lu, W.Q., Sun, B., Wang, Q.,
741 Zhang, B., Chavarro, J.E., Zhang, J., Pan, A., 2019b. Seminal plasma metabolome in relation
742 to semen quality and urinary phthalate metabolites among Chinese adult men. Environ. Int. 129,
743 354-363.
744 Wild, C.P., 2005. Complementing the genome with an "exposome": The outstanding challenge
745 of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol.
746 Biomarkers Prev. 14 (8), 1847-1850.
747 Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vazquez-Fresno, R., Sajed, T.,
748 Johnson, D., Li, C.R., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S.,

749 Arndt, D., Liang, Y.J., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y.F., Mandal, R., Neveu,
750 V., Pon, A., Knox, C., Wilson, M., Manach, C., Scalbert, A., 2018. HMDB 4.0: the human
751 metabolome database for 2018. Nucleic Acids Res. 46 (D1), D608-D617.
752 Wu, Y., Ding, R., Zhang, X., Zhang, J., Huang, Q., Liu, L., Shen, H., 2021. Meet-in-metabolite
753 analysis: A novel strategy to identify connections between arsenic exposure and male infertility.
754 Environ Int 147, 106360.
755 Wu, Y., Zhang, J., Peng, S.Y., Wang, X.F., Luo, L.Z., Liu, L.P., Huang, Q.Y., Tian, M.P., Zhang,
756 X.Q., Shen, H.Q., 2018. Multiple elements related to metabolic markers in the context of
757 gestational diabetes mellitus in meconium. Environ. Int. 121, 1227-1234.
758 Yan, Z.X., Yan, R., 2015. Improved Data-Dependent Acquisition for Untargeted Metabolomics
759 Using Gas-Phase Fractionation with Staggered Mass Range. Anal. Chem. 87 (5), 2861-2868.
760 Yang, L.L., Wang, S., Peng, X., Zheng, M.H., Yang, Y.P., Xiao, K., Liu, G.R., 2019. Gas
761 chromatography-Orbitrap mass spectrometry screening of organic chemicals in fly ash samples
762 from industrial sources and implications for understanding the formation mechanisms of
763 unintentional persistent organic pollutants. Sci. Total Environ. 664, 107-115.
764 Yao, Y.Y., Wang, P.G., Shao, G., Del Toro, L.V.A., Codero, J., Giese, R.W., 2016. Nontargeted
765 analysis of the urine nonpolar sulfateome: a pathway to the nonpolar xenobiotic exposome.
766 Rapid Commun. Mass Spectrom. 30 (21), 2341-2350.
767 Yeung, L.W.Y., Mabury, S.A., 2016. Are humans exposed to increasing amounts of unidentified
768 organofluorine? Environ. Chem. 13 (1), 102-110.
769 Zhang, J., Huang, Z., Chen, M., Xia, Y., Martin, F.L., Hang, W., Shen, H., 2014a. Urinary
770 metabolome identifies signatures of oligozoospermic infertile men. Fertil Steril 102 (1), 44-53

e12.

Zhang, J., Mu, X., Xia, Y., Martin, F.L., Hang, W., Liu, L., Tian, M., Huang, Q., Shen, H., 2014b. Metabolomic analysis reveals a unique urinary pattern in normozoospermic infertile men. J Proteome Res 13 (6), 3088-3099.

Zhang, J., Shen, H., Xu, W., Xia, Y., Barr, D.B., Mu, X., Wang, X., Liu, L., Huang, Q., Tian, M., 2014c. Urinary metabolomics revealed arsenic internal dose-related metabolic alterations: a proof-of-concept study in a Chinese male cohort. Environ Sci Technol 48 (20), 12265-12274.

Zheng, X.Y., Wojcik, R., Zhang, X., Ibrahim, Y.M., Burnum-Johnson, K.E., Orton, D.J., Monroe, M.E., Moore, R.J., Smith, R.D., Baker, E.S., 2017. Coupling Front-End Separations, Ion Mobility Spectrometry, and Mass Spectrometry For Enhanced Multidimensional Biological and Environmental Analyses, in: Cooks, R.G., Pemberton, J.E. (Eds.), Annual Review of Analytical Chemistry, Vol 10. Annual Reviews, Palo Alto, pp. 71-92.

Zhou, Z.W., Tu, J., Zhu, Z.J., 2018. Advancing the large-scale CCS database for metabolomics and lipidomics at the machine-learning era. Curr. Opin. Chem. Biol. 42, 34-41.