# Utilizing a biology-driven approach to map the exposome in health and disease

Chung, Ming Kei; Rappaport, Stephen M.; Wheelock, Craig E.; Nguyen, Vy Kim; van der Meer, Thomas P.; Miller, Gary W.; Vermeulen, Roel; Patel, Chirag J.

*Published in:*
Environmental Health Perspectives

*DOI:*
10.1289/EHP8327

A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP8327.

# Utilizing a Biology-Driven Approach to Map the Exposome in Health and Disease: An Essential Investment to Drive the Next Generation of Environmental Discovery

*Ming Kei Chung,[1] Stephen M. Rappaport,[2] Craig E. Wheelock,[3] Vy Kim Nguyen,[4,5] Thomas P. van der Meer,[6] Gary W. Miller,[7] Roel Vermeulen,[8] and Chirag J. Patel[1]*

[1]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA
[2]Program in Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, California, USA
[3]Division of Physiological Chemistry 2, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden
[4]Department of Environmental Health Sciences, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA
[5]Department of Computational Medicine and Bioinformatics, Medical School, University of Michigan, Ann Arbor, Michigan, USA
[6]Department of Endocrinology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[7]Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, USA
[8]Utrecht University & Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

**Background:** Recent developments in technologies have offered opportunities to measure the exposome with unprecedented accuracy and scale. However, because most investigations have targeted only a few exposures at a time, it is hypothesized that the majority of the environmental determinants of chronic diseases remain unknown.

**Objectives:** We describe a functional exposome concept and explain how it can leverage existing bioassays and high-resolution mass spectrometry for exploratory study. We discuss how such an approach can address well-known barriers to interpret exposures and present a vision of next-generation exposomics.

**Discussion:** The exposome is vast. Instead of trying to capture all exposures, we can reduce the complexity by measuring the functional exposome—the totality of the biologically active exposures relevant to disease development—through coupling biochemical receptor-binding assays with affinity purification–mass spectrometry. We claim the idea of capturing exposures with functional biomolecules opens new opportunities to solve critical problems in exposomics, including low-dose detection, unknown annotations, and complex mixtures of exposures. Although novel, biology-based measurement can make use of the existing data processing and bioinformatics pipelines. The functional exposome concept also complements conventional targeted and untargeted approaches for understanding exposure-disease relationships.

**Conclusions:** Although measurement technology has advanced, critical technological, analytical, and inferential barriers impede the detection of many environmental exposures relevant to chronic-disease etiology. Through biology-driven exposomics, it is possible to simultaneously scale up discovery of these causal environmental factors. https://doi.org/10.1289/EHP8327

## Introduction

### *Environment in Health and Disease*

Genes, environmental factors, and their interaction are thought to contribute to the development of chronic diseases. Extensive analyses from twin and health-record studies have revealed that—except for a handful of rare disorders—genetics on average explains 30–50% of disease risk (Czene et al. 2002; Lakhani et al. 2019; Polubriaginof et al. 2018). On the other hand, common environmental factors—including diet, smoking, and air pollution—together may contribute about 46% of global deaths (Lim et al. 2012; Rappaport 2016). The specific environmental drivers remain elusive for many complex chronic diseases, and it is difficult to disentangle the interactions among the elements of mixture exposures (Chung et al. 2018; Patel and Ioannidis 2014; Patel and Manrai 2015).

Address correspondence to Ming Kei (Jake) Chung, Department of Biomedical Informatics, Harvard Medical School, Harvard University, 10 Shattuck St., Room 302, Boston, MA 02115. Email: jake_chung@hms.harvard.edu or jakemk@berkeley.edu, or Chirag J. Patel, Department of Biomedical Informatics, Harvard Medical School, Harvard University, 10 Shattuck St., Room 302, Boston, MA 02115. Telephone: (617) 432-1195. Email: chirag_patel@hms.harvard.edu

The authors declare they have no actual or potential competing financial interests.

Received 22 September 2020; Revised 28 March 2021; Accepted 13 July 2021; Published 26 August 2021.



The exposome encompasses the life course of exposures received by an individual (Wild 2005) and provides a unified framework for discovering environmental determinants of diseases (Rappaport and Smith 2010). The exposome has been embraced by both governmental (Birnbaum 2012) and research communities (Lippmann 2013; Wheelock and Rappaport 2020) and has motivated initiatives focusing on early life exposures and life course approaches investigating the impacts of complex environmental exposures (Vineis et al. 2017; Vrijheid et al. 2014). It has also stimulated development of new technologies to study exposures based upon measurements obtained from both inside (Dennis et al. 2016) and outside of the body (Jiang et al. 2018) using omics, remote sensing, and portable sensors (Chung and Patel 2019; Vineis et al. 2017). Data derived from such investigations not only characterize the exposome but also offer clues on potentially causal exposures. In the following, we define "environmental exposure" as an exogenous factor that interacts physically with its human host, and a "biologically-active" exposure as a molecular exposure capable of inducing a biological change at a detectable level in the human population.

### *Data-Driven Exposomics*

Almost all epidemiological studies of environmental factors are hypothesis driven. That is, they target selected pollutants or dietary and lifestyle factors to identify associations with a particular disease to quantify exposure–response relationships. In contrast, data-driven studies are being increasingly applied to discover risk factors for disease, particularly for genetic epidemiology. For example, genome-wide association studies (GWASs) employ a top-down strategy to simultaneously investigate associations between millions of genetic variants and a disease outcome (Tam et al. 2019). To establish an equivalent top-down strategy for environmental epidemiology, exposome-wide association studies (EWASs) are being

designed to simultaneously detect many thousands of small-molecule features in biospecimens and compare their abundances between healthy and diseased subjects, preferably from a longitudinal cohort (Rappaport 2012, 2018; Rappaport and Smith 2010; Manrai et al. 2017). The approach is hypothesized to overcome challenges in multiplicity and fragmented literature of associations that occur due to nonsystematic testing (Ioannidis et al. 2009; Patel and Ioannidis 2014; Manrai et al. 2019).

Researchers can seek out environmental risk or protective factors with broad-scale targeted measurement (Balik-Meisner et al. 2018; Patel et al. 2010; Sipes et al. 2013) or screen for relevant exposures with semitargeted approaches, such as suspect screening, which search for a predefined and typically by-class chemical space (Andra et al. 2017; Wang et al. 2018). Currently, untargeted measurement of the blood exposome is conducted with high-resolution mass spectrometry (HRMS) because of its capability to distinguish peaks in mass spectra with high mass accuracy (i.e., at the level of <1 ppm). Several studies have demonstrated the capability of HRMS to detect a large number of both endogenous and exogenous exposures in observational studies (Board on Life Sciences; Division on Earth and Life Studies; National Academies of Sciences, Engineering, and Medicine 2016; Go et al. 2015a; Jamin et al. 2014; Roca et al. 2014). Although HRMS has largely been used to discover changes in biological pathways associated with known exposures (Turner et al. 2018; Vrijheid et al. 2014; Walker et al. 2016b; Yuan et al. 2016) or to find biomarkers of diseases (Lu et al. 2016; Osborn et al. 2013; Roede et al. 2013; Wang et al. 2011), there is now an impetus to apply this technology to identify environmental risk factors in biomedical and epidemiological studies (Rappaport 2018).

### *Challenges of Data-Driven Exposomics in Observational Human Research*

We describe two main challenges for wide application of data-driven exposomics. First, the environmental exposures associated with chronic diseases are likely to be low dose. Summarizing from two databases, concentrations of many exogenous compounds in blood—such as dietary micronutrients, phytoestrogens, microbial metabolites, drugs, consumer chemicals, and pollutants—are <5 μM and have a high range of limit of detection (Rappaport et al. 2014). One of the major distinctions between targeted and untargeted mass spectrometry (MS) is the tradeoff between sensitivity and coverage; targeted MS involves optimized methodology for selected compounds for better sensitivity and specificity. This means that a sizable fraction of the exposome may not be detected by untargeted HRMS. Second, the annotation of features is often incomplete in HRMS. During analysis, a large proportion of the signals are uninterpretable because they are the "dark matter" of the exposome—that is, they lack annotation or have ambiguous identity (Vermeulen et al. 2020). Putative annotation is generally performed by matching either or both of the accurate mass and spectrum of the features to the known records in reference databases, which can be specialized for endogenous metabolites (Kanehisa and Goto 2000; Wishart et al. 2018) or xenobiotic compounds (Meijer et al. 2021; Neveu et al. 2017; Wishart et al. 2015, 2008). This can contribute to reporting bias because xenobiotic databases contains only ∼1,000–70,000 chemicals. In contrast, PubChem contains >90 million records (Kim et al. 2019), but only a small fraction of this database comes with information related to sources and biological activities. This puts a premium on follow-up investigations to validate potentially causative features with independent biospecimens and chemically identify the most promising hits.

So far, the rapid development in exposure assessment technology and infrastructure has made the exposome concept more concrete. While significant progress has been made to advance our understanding on how candidate exposures could influence human health, we still lack the capacity for answering one of the most fundamental questions in biomedical research: what environmental factors, in aggregate and in their totality, that can cause or prevent a disease?

## Discussion

### *Biology-Driven Exposomics*

We are likely decades away from measuring the life course exposome of an individual. A top-down strategy provides a way to explore the blood exposome, but challenges from miniscule and complex exposures have become a major hurdle in data-driven exposomics. We hypothesize that, if the ultimate goal is to find the environmental causes of diseases, we do not need to measure all the entities. We believe that a pragmatic view is to focus on those chemicals that *a*) we are exposed to and *b*) are biologically active—that is, capable of inducing a biological change at the detected level in the population. From a molecular view, almost all exposures, including chemicals entering the body and initial signaling molecules synthesized by the body due to nonchemical stimuli (e.g., mental stress and light), exert their biological effects through interacting with functional biomolecules (Peters et al. 2021; Vermeulen et al. 2020), such as DNA, mRNA (riboswitches), and proteins (receptors and enzymes).

We define the functional exposome as encompassing the life course exposures to both endogenous and exogenous biologically active molecules, which include parent compounds and metabolites from host and microbes (Chen et al. 2019), enantiomers (Fischer et al. 2014), or reactive forms of chemicals (Chung et al. 2013; Smela et al. 2002). The functional exposome contains a reduced set of exposures at the interface of environment (nurture) and biology (nature) that is important for explaining the variation in disease among individuals. Indeed, the essence of the functional exposome has been suggested by early exposome advocates (Miller and Jones 2014; Rappaport and Smith 2010; Wild 2012), and finding the determinants of health and diseases through the functional exposome is analogous to using expressed sequence tags to speed up gene discovery in the early days of genomics (Adams et al. 1991). We define biology-based measurement (BBM; Figure 1) as the collection of analytical methods or technologies that can characterize the molecular interaction between an exposure and a functional biomolecule and thereby perform functional exposome measurement (FEM). Functional exposomics aims to identify the functions and sources of biologically active exposures. It can be broadly divided into 3 steps: 1) measuring a biologically active subset of the exposome; 2) charactering the connections between exposures and diseases; 3) identifying the transformation and sources of the exposures.

### *Types of Molecular Binding*

Broadly speaking, molecular binding can be classified into two types (Klaassen 2018). First, some interactions result in irreversible nonspecific binding of reactive electrophiles to functional macromolecules. These reactions form adducts that can change the biological activity of the biomolecules or initiate mutations that can induce cancers. Second, another class of interactions includes reversible (noncovalent) and specific binding caused by ligands, which can be small molecules, ions, peptides, or proteins that modulate the biological function of a target.

### *Approaches to Search for Biologically Active Exposures*

Effect-directed analysis shares the same goal of BBM and is an effective yet tedious approach for finding compounds of interest

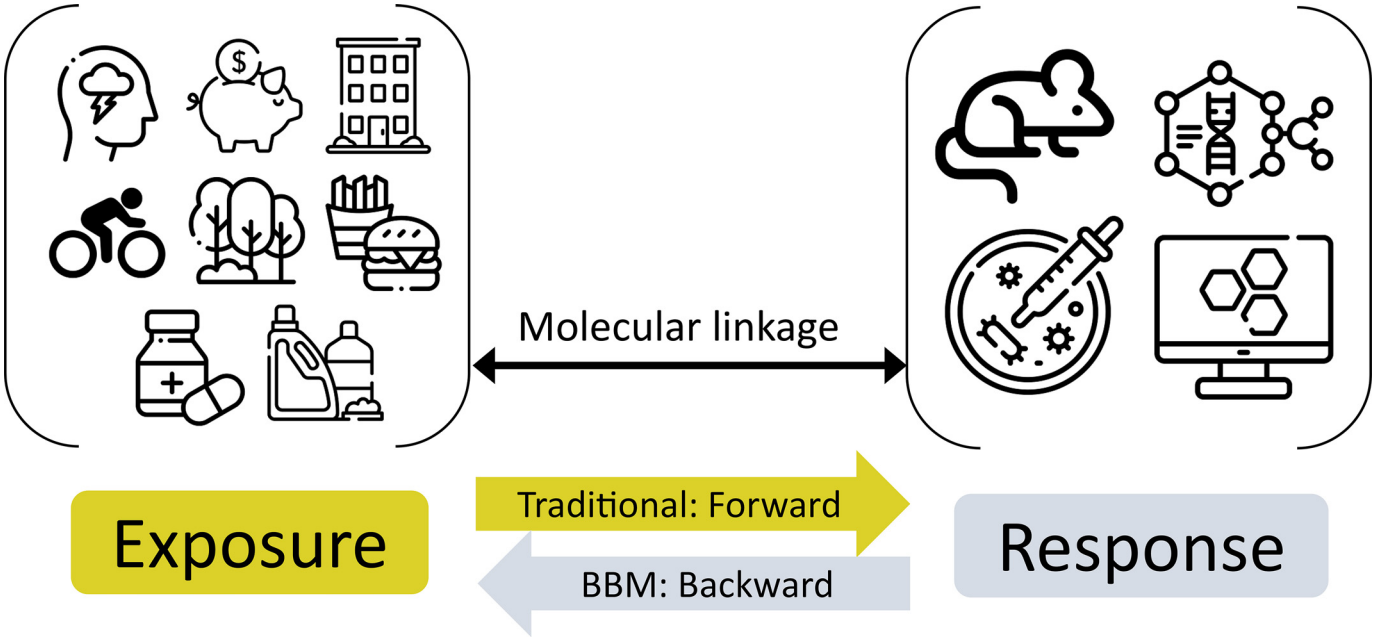


**Figure 1.** Approaches to studying the exposure–response relationship. The exposome can be broadly divided into four distinct domains: living environment, socioeconomic factors, lifestyle, and physical-chemical exposures. The effects of exposures can be studied at the ecological, *in vivo*, *in vitro*, molecular, and *in silico* levels. Although exposure and response are heterogeneous, almost all causal biological effects from an exposure are explainable by the underlying molecular connections. To investigate a relationship, traditional study takes a forward approach—from exposure to response. Examples include a hypothesis-driven design with a few predefined exposures and a semi-/fully data-driven design to conduct exploratory studies. In contrast, studies using a biology-based measurement (BBM) approach work differently—from response to exposure—at the molecular level. The bait to capture the exposures can be a few predefined, a particular class, or a broad-scale selection of functional biomolecules. Certain icons in the figure were made by Freepik (www.flaticon.com).

in a mixture (Brack et al. 2016). It involves an iterative sample fractionation process by chromatography, with effect assessment spanning from molecular bindings to whole-organism assays. This approach has been successfully applied to identify an ozonation product of the common tire chemical *N*-(1,3-dimethylbutyl)-*N*′-phenyl-*p*-phenylenediamine (6PPD) called 6PPD-quinone in stormwater. For decades, 6PPD-quinone has been responsible for the sudden death of coho salmon (Tian et al. 2021).

Because reactive chemicals have short lifespans *in vivo*, their adducts are used as proxies for exposures (Chung et al. 2010). A chemoproteomic platform has been developed to conduct proteome-wide mapping of targeted environmental electrophiles (Medina-Cleghorn et al. 2015), and untargeted adductomics has also demonstrated its potential to discover exposures to reactive electrophiles (Carlsson et al. 2019; Grigoryan et al. 2016; Rappaport et al. 2012).

For studying reversible binding, affinity selection–MS is one option. Briefly, a selection of proteins or enzymes is incubated with the mixtures and the bound complexes are separated with size exclusion chromatography and characterization of the active compounds by MS (Annis et al. 2007; Yang et al. 2020). Affinity purification–MS (AP-MS) is a similar technology; however, active compounds in the mixtures are pulled down by immobilized biomolecules (Dunham et al. 2012).

Affinity-based methods have been applied to proteomics and molecular biology to understand the interaction and activity of molecules (Morris et al. 2014; Rinschen et al. 2019). Yet, there has been no known attempt to apply an untargeted platform for systematic identification of biologically active molecules in a human observational study. Given that receptor binding may be relevant to the development of complex chronic diseases, we focus our remaining discussion on ligand–receptor binding and the adaptation of the binding assay into the affinity purification format, which is one of the BBM approaches for FEM. We then discuss how BBM, exemplified by AP-MS, can address well-known bottlenecks in data-driven studies and our view on next-generation exposomics.

### *Binding Assays to Assess Effects of Exposures*

Traditional molecular studies typically characterize the receptor activity of known ligands using either binding or functional reporter assays (Seethala and Zhang 2016). The former is a type of biochemical-based assay using a purified recombinant protein and engineered ligands, which are radiolabeled or fluorophore-tagged for generating signals proportional to the binding level. The latter is cell based and uses genetically modified cells to produce a detectable signal upon receptor activation. These assays are also used in *in vitro* high-throughput screening for drug discovery in the pharmaceutical industry with >1 million compounds (Carnero 2006; Mayr and Bojanic 2009) and for compound toxicity in the U.S. Environmental Protection Agency's ToxCast project with >1,858 chemicals and 821 assays (Filer et al. 2014; Sipes et al. 2013).

### *Measuring the Functional Exposome*

Above, we defined the functional exposome as the subset of the exposome that consists of biologically-active exposures present in an individual. When both ligand and receptor are known, an optical system is a simple and sensitive choice for measurement (Seethala and Zhang 2016). In the context of data-driven exposomics, exposures are unknowns that require structural characterization and thus cannot be detected by existing binding assays. However, biochemistry-based assays can be adapted to screen for classes of ligands with AP-MS (Dunham et al. 2012). For example, tagged recombinant receptors (bait) can be immobilized on a microarray that bind to specific ligands (prey) during incubation. After removing the unbound sample, the intensity and mass spectra of the recovered ligands (putative exposures that were bound

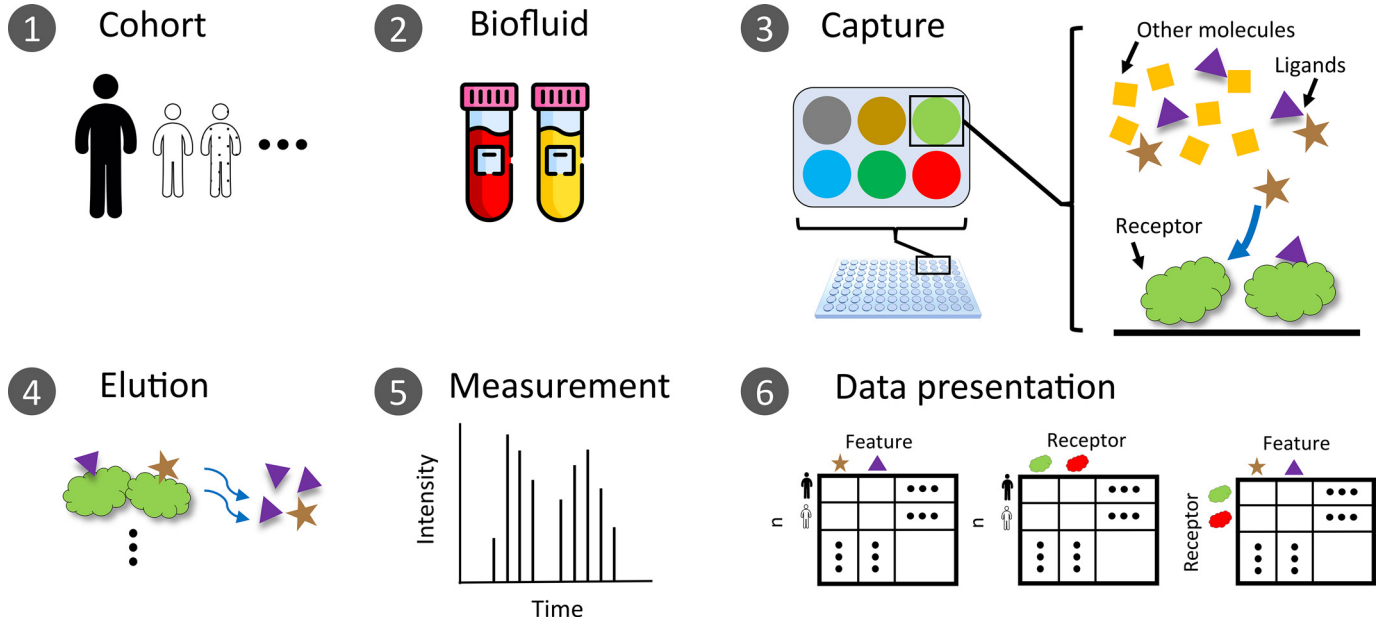


**Figure 2.** Using a biology-based measurement approach to characterize the functional exposome. The assay can be implemented in the form of affinity purification–mass spectrometry in six steps. (1) A group of individuals (*n*) is sampled. (2) Biofluids, such as plasma, are isolated from blood and (3) loaded to the microplate (each well is coated with one type of receptor) to capture circulating ligands. (4) After incubation, ligands are extracted and (5) characterized by mass spectrometry. (6) Data can be presented in three different matrices (*n* by feature, *n* by receptor, and receptor by feature) with signals integrating across features or receptors for downstream analyses. A key advantage of this approach is that functional analysis of the exposome is possible without upfront feature identification because molecular exposures are biophysically linked to receptors with known functions. Certain icons in the figure were made by Freepik (www.flaticon.com).

to a receptor) are measured with HRMS (Figure 2). Here, we describe only one of the many approaches that leverages existing methodologies for BBM. The workflow is modular and substitutable. For example, baits can be immobilized on microbeads and ligands can be detected by various HRMS platforms (Lanucara et al. 2014; Morris et al. 2014; Wieser et al. 2012).

The entire data preprocessing pipeline for HRMS can be applied to the raw data generated by AP-MS, including feature detection, filtering, alignment, and normalization (Dunn et al. 2011). There are online resources (Tautenhahn et al. 2012; Xia et al. 2009) and offline software (Uppal et al. 2013; Yu et al. 2009) for data and bioinformatic processing, and data can be exported for additional visualization and statistical analyses by, for example, a correlation globe (Chung et al. 2018; Patel and Manrai 2015). Depending on the research questions, the signals can be summarized and presented by three matrices: subject by feature, subject by receptor, and receptor by feature (Figure 2). The first data representation will fit directly into an EWAS framework in which chemical features are used as the input variables.

### *Opportunity and New Areas*

Although BBM quantifies chemical exposures akin to current-day HRMS, it has three innovative aspects that provide solutions to solve some of the well-known barriers in data-driven exposomics in real-world scenarios (Table 1; Figure 3A–C). Specifically, the challenges include low dose exposures, annotation of features in untargeted measurement, and mixture exposures.

***Enrichment and high-throughput screening.*** Much of the blood exposome consists of low-abundance features that are poorly characterized by untargeted HRMS, sacrificing sensitivity for wider coverage (Rappaport et al. 2014). The pull-down format of AP-MS is equivalent to enriching the samples—keeping potentially interesting ligands while removing the nonbinding species—thereby increasing the signal-to-noise ratio and sensitivity of the measurement (Chung et al. 2014; Florentinus-Mefailoski et al. 2014).

Another benefit of AP-MS is the optional use of liquid chromatography (LC) to separate complex components. Unlike many

**Table 1.** Expected opportunities and limitations of using a biology-based measurement approach in exposomics studies.

| Opportunities | Limitations and uncertainties |
|---|---|
| The functional exposome is mapped to the genome by biophysical interactions in the data generation step, providing mechanistic insights for statistically associated relationships. | Not all receptors can be directly used without further assay development (e.g., transmembrane receptors). |
| The workflow is modular and capable of switching to different types of functional biomolecules (e.g., DNA, RNA, protein) and different omics platforms to measure the ligands (e.g., proteomics for peptide ligands). | Compared with the gold standard of measurement (LC-MS/MS), detection rate in the general population is still an unknown. |
| The approach leverages existing biochemical and MS technologies. Minimal development is needed. | Exact biofluid volume requirement per person is uncertain but, based on experience, is likely in the range of 300–600 μL. |
| Approach is tailored to discovery study in exposomics; data is used to drive EWASs for finding new exposures associating with diseases. | Although upfront compound confirmation of the unknown is not required, ultimately, structural characterization is still needed to link sources and additional information from the literature. |
| Environmental samples, such as dust extracts and water samples, can complement biological samples to provide a bigger picture of the functional exposome. | Non-host compounds are expected to be dominating in the biofluid measurement but this awaits characterization and confirmation. |

Note: LC, liquid chromatography.

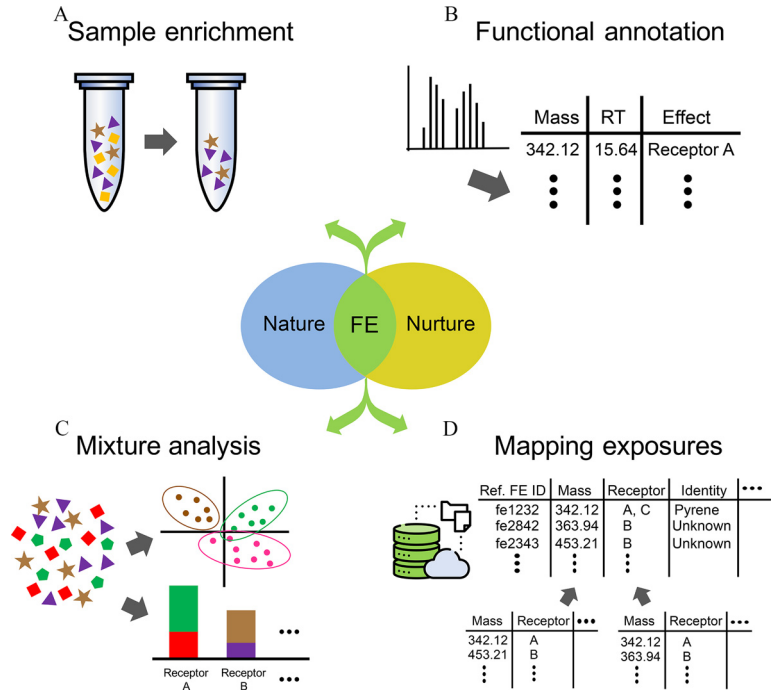


**Figure 3.** Key advantages of using a biology-based measurement approach to study the functional exposome (FE), where nature meets nurture. (A) Enrichment is not required as an extra step but, rather, intrinsically provided by the pull-down assay format. (B) Each detected feature is functionally annotated by the corresponding binding receptors. (C) Health impacts can be assessed in two ways: coexposure correlation or shared mode of action. (D) Each feature can be assigned to a unique reference FE identifier (Ref. FE ID), which is characterized by an accurate mass, the functional molecules (e.g., binding receptors), and other attributes (e.g., mass spectra, detection frequency in the population). The mapping, cataloging, and sharing of exposure information allow unambiguous communication of knowns and unknowns across studies. Certain icons in the figure were made by Freepik (www.flaticon.com). Note: RT, retention time.

optical detection systems capable of massive parallel measurement, LC-HRMS operates in a serial fashion, and a chromatography gradient profile typically takes 30–60 min per sample to produce (De Vos et al. 2007; Yuan et al. 2012). Enriched samples have a clear matrix and therefore can be analyzed directly by novel commercialized high-throughput HRMS, taking only 10–60 s per sample (Kempa et al. 2019).

***Annotating the dark matter.*** The Metabolomics Standards Initiative has set identification standards of features from MS that have four levels of confidence (Sumner et al. 2007): *a*) unknown compounds that are differentiable and quantified; *b*) putative characterization by compound class; *c*) putative identification by matching either or both of the accurate mass and spectrum to the records in libraries; and *d*) compound confirmation by comparing properties with those of an authentic reference standard. Untargeted HRMS can detect 10,000–20,000 features in blood samples, but even though there are accurate mass measures and retention times, up to 80% cannot be identified (Uppal et al. 2016), and of those putatively identified, only a small fraction can be reported with higher confidence (Go et al. 2015b; Walker et al. 2016a, 2016b). Further, sources and biological effects of the identified compounds may not be available in the literature, making meaningful inference of their potential roles in disease development difficult. This is an important drawback for using conventional untargeted HRMS to investigate the biological plausibility of potential health effects arising from thousands of features.

In contrast, each feature detected by BBM is a molecular entity that is physically bound to a receptor modulating a known biological function and is, therefore, intrinsically interesting and functionally interpretable without *a priori* feature identification. To accelerate the pace of discovery, functional annotation complements existing approaches to eventually identify unknown features. For example, binding information could enhance the interpretability of correlation-based network analysis of features (Quinn et al. 2017; Uppal et al. 2016).

***Analyzing mixtures.*** Humans are exposed to complex chemical mixtures from endogenous and exogenous sources that are common to everyday life (Rappaport and Smith 2010). Past epidemiological studies have tended to analyze only one or a few exposures and confounders at a time. This practice can inflate false positives and lead to publication bias because the effects of multiple comparisons are typically not considered (Ioannidis 2005, 2008) or may be confounded by variables not assayed in the study. In addition, it eliminates the opportunity to study how multiple exposures can interact in disease processes. In its 2018–2023 strategic plan, the National Institute of Environmental Health Sciences highlighted this problem as a focus area for future research (Birnbaum 2017).

To address this problem, we can analyze multiple exposures together with statistical methods that are robust to model assumptions or account for the correlation among predictors. These methods generally involve data reduction to enhance interpretation—methods such as (sparse) partial least squares regression, elastic net, and Bayesian kernel machine regression (Agier et al. 2016; Lazarevic et al. 2019; Taylor et al. 2016)—and have been compared in terms of computational performance and strength and weakness—such as bias-variance tradeoff and in epidemiological settings (James et al. 2017; Patel 2017)—to understand the best-use scenarios.

In toxicology, the mixture problem is addressed with an effect-based approach. Exposures with structural similarity (e.g., polychlorinated biphenyls and dioxins; polycyclic aromatic hydrocarbons; xenoestrogens) generally act together because they share the same mode of action (i.e., acting on the same receptor). The joint effects can reasonably be assumed to be

additive, and hence their effects can be expressed in a relative scale as toxic equivalency factors (TEFs). The total toxicity of a mixture, called the toxic equivalent, is the weighted sum of the individual TEFs by concentration (Nisbet and LaGoy 1992; Van den Berg M et al. 1998).

BBM presents mixture exposure data in two ways (Figure 2). For features as input variables (i.e., integrating intensity across receptors), we can apply off-the-shelf regression approaches to identify associations and coexposure patterns. On the other hand, a complex mixture contains thousands of compounds and those with the same mode of action mostly have additive effects (Kortenkamp 2014). The rest of the compounds act independently because synergy generally does not occur at low doses (Escher et al. 2020). Assuming receptor activation is proportional to ligand binding affinity, the overall effect of a mixture on a receptor can be estimated by integrating the intensity across features.

### *Envisioning the Next-Generation Exposomics*

From another perspective, BBM is essentially an approach to map the exposome to biological function. This mapping strengthens the functional analysis of exposures because biological inference is both statistical and mechanistic—in this case, involving biophysical interactions with receptors of known biological function.

HRMS generates accurate mass for library matching. Even so, multiple compound matches for a single feature is common and the ambiguity makes sharing the masses of features hardly useful (Uppal et al. 2016). Because binding is structurally specific to each receptor, in addition to mass and spectrum, BBM can use mapping to receptors as an orthogonal parameter suitable for assigning unique functional exposome identifiers (similar to reference single nucleotide polymorphism identifiers in GWASs). We claim that creating a catalog of distinct and nonredundant entries of known chemicals and unknown features in a central sharing platform opens new ways to interpret the exposures and to permit systematic comparisons across studies (Figure 3D).

### *Crowdsourcing the Unknown Exposures in the Catalog*

The exposome is a growing entity. Since 1800, new chemical compounds have been synthesized at a stable 4.4% annual growth rate (Llanos et al. 2019). Still, it is possible to gain additional insights on the unknown exposures with an open-access and crowdsourced catalog. First, unknown exposures could be systematically analyzed across multiple diseases. Assume exposure *A* binds to receptor *K*. If *A* was associated with two diseases in the same category (e.g., both are neurological), then the exposure–receptor pair could be mechanistically related to disease development. The exposures with broader biological significance could be priority exposures, whereas those with high specificity to a particular disease could play a causal role. Second, it is possible to glean the spatial patterns and temporal variability of the exposures from multiple studies and establish a background level for comparison (Vermeulen et al. 2020). We can ask questions such as the following: Is it a long term exposure? Is there seasonality? Does the prevalence of an unknown exposure match that of a disease? Reproducibility of the association between an unknown and a disease could also be assessed. Third, the research community can coordinate their resources to conduct structural identification of unknown features. Priority and importance of features can be assigned by voting, and the catalog can be updated with the latest identification results. In short, the catalog could reinforce causal inference (Ioannidis 2016) and coalesce knowledge gained on the roles of known and unknown exposures from field-specific studies.

### *Fine Mapping of Exposures to Mixtures*

We foresee that in the coming decade, conventional targeted and untargeted measurement approaches will continue to be harnessed in observational studies because they will still be cheaper than BBM. Nevertheless, conventional studies could, we predict, indirectly benefit from BBM when the significant signals or targeted interests overlap with the records in the catalog.

Fine mapping of associated or selected exposures to meaningful mixtures may have three implications. First, it may provide biological interpretation of the statistical associations of EWASs. For example, are the exposures associated with a disease mostly the ligands of a particular receptor? Second, the catalog may guide the design of hypothesis-driven or experimental mixture study. Conventional targeted studies test a predefined exposure–disease relationship. By using the predefined exposure as a marker, the catalog may provide a set of meaningful mixtures for further investigation. Third, the mapping information may facilitate rare exposure analysis. Most of the EWASs include a data filtering step to remove rare exposures prior to statistical analysis. Assuming the exposures binding to the same receptor or same category of receptors have similar biological effects, their signals could be combined and thus increase the statistical power to detect a potential effect (Chung et al. 2019).

### *Technology Development*

There are many different functional protein targets for BBM, such as plasma proteins, transport proteins, chemoreceptors, and enzymes. At present, there are 48 nuclear receptors (Mangelsdorf et al. 1995) and roughly 750 G protein-coupled receptors (Vassilatis et al. 2003). A small subset of receptors (e.g., estrogen and androgen receptors) relevant to some of the ubiquitous synthetic or natural xenoestrogens such as genistein, phthalates, and bisphenol A could be selected (Lee et al. 2013) (Table 1). In the future, successful development of a bio-bait microarray will lower the requirements for large-scale and routine profiling of the functional exposome, such as in the National Health and Nutrition Examination Survey (Sobus et al. 2015). Although the measurement technologies for BBM can be borrowed from other fields, significant investment and concerted efforts are still needed for optimizing sensitivity, scaling up capacity, setting up data collection and sharing standards, and assessing reproducibility.

## Conclusions

We acknowledge the monumental advancements made in exposomics in the past 10 years that focused largely on answering the "how" questions, such as how to get more accurate and convenient measurements and how a molecular or complex exposures can exert effects on the development of chronic diseases. In contrast, biology-driven exposomics focuses on the "what" questions to help scale the search for causal exposures. In this commentary, we have articulate conceptually how AP-MS can address well-known barriers in data-driven studies. Creating a biophysical map of the exposome to guide mixture studies will help to lower the barrier of entry for exposome research between disciplines in public health and biomedical sciences.

In the era of next-generation exposomics, we envision that health researchers will have accessible tools that strike a balance between answering questions of how and what. In our view, through tighter integration with the genome and other subcellular networks, we will have better capability to study the totality of exposures and decipher the known unknowns and unknown unknowns in the exposome. We need to invest and embrace a new technological approach that can bring us closer to identifying the environmental causes of complex diseases.

## Acknowledgments

This study was supported by grants from the U.S. National Institutes of Health (NIH) through the National Institute for Environmental Health Sciences (S.M.R.: P01ES018172, P50ES018172, and P42ES004705; V.K.N.: R01ES028802), the Swedish Heart-Lung Foundation (C.E.W.: Hjärt-Lungfonden 20170736 and 20180290), the NIH (G.W.M.: U2CES030163 and RC2 DK118619), the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research through the EXPOSOME-NL (R.V.; 024.004.017), the European Union's Horizon 2020 research and innovation programme (R.V.: 874627), and the NIH through the National Institute for Allergy and Infectious Diseases (C.J.P.: AI127250) and the National Institute of Mental Health and National Institute of Allergy and Infectious Diseases (M.K.C: R01MH107205 and R01AI127250).

## References

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, et al. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science 252(5013):1651–1656, PMID: 2047873, https://doi.org/10.1126/science.2047873.

Agier L, Portengen L, Chadeau-Hyam M, Basagaña X, Giorgis-Allemand L, Siroux V, et al. 2016. A systematic comparison of linear regression–based statistical methods to assess exposome-health associations. Environ Health Perspect 124(12):1848–1856, PMID: 27219331, https://doi.org/10.1289/EHP172.

Andra SS, Austin C, Patel D, Dolios G, Awawda M, Arora M. 2017. Trends in the application of high-resolution mass spectrometry for human biomonitoring: an analytical primer to studying the environmental chemical space of the human exposome. Environ Int 100:32–61, PMID: 28062070, https://doi.org/10.1016/j.envint.2016.11.026.

Annis DA, Nickbarg E, Yang X, Ziebell MR, Whitehurst CE. 2007. Affinity selection-mass spectrometry screening techniques for small molecule drug discovery. Curr Opin Chem Biol 11(5):518–526, PMID: 17931956, https://doi.org/10.1016/j.cbpa.2007.07.011.

Balik-Meisner M, Truong L, Scholl EH, La Du JK, Tanguay RL, Reif DM. 2018. Elucidating gene-by-environment interactions associated with differential susceptibility to chemical exposure. Environ Health Perspect 126(6):067010, PMID: 29968567, https://doi.org/10.1289/EHP2662.

Birnbaum LS. 2012. NIEHS's new strategic plan. Environ Health Perspect 120(8): a298, PMID: 22853936, https://doi.org/10.1289/ehp.1205642.

Birnbaum LS. 2017. Updating the NIEHS strategic plan. Environ Health Perspect 125(7):071001, PMID: 28749368, https://doi.org/10.1289/EHP2502.

Board on Life Sciences; Division on Earth and Life Studies; National Academies of Sciences, Engineering, and Medicine. 2016. *Use of Metabolomics to Advance Research on Environmental Exposures and the Human Exposome: Workshop in Brief*. Washington, DC: National Academies Press.

Brack W, Ait-Aissa S, Burgess RM, Busch W, Creusot N, Di Paolo C, et al. 2016. Effect-directed analysis supporting monitoring of aquatic environments—an in-depth overview. Sci Total Environ 544:1073–1118, PMID: 26779957, https://doi.org/10.1016/j.scitotenv.2015.11.102.

Carlsson H, Rappaport SM, Törnqvist M. 2019. Protein adductomics: methodologies for untargeted screening of adducts to serum albumin and hemoglobin in human blood samples. High Throughput 8(1):6, PMID: 30857166, https://doi.org/10.3390/ht8010006.

Carnero A. 2006. High throughput screening in drug discovery. Clin Transl Oncol 8(7):482–490, PMID: 16870538, https://doi.org/10.1007/s12094-006-0048-2.

Chen S, Henderson A, Petriello MC, Romano KA, Gearing M, Miao J, et al. 2019. Trimethylamine N-oxide binds and activates PERK to promote metabolic dysfunction. Cell Metab 30(6):1141–1151.e5, PMID: 31543404, https://doi.org/10.1016/j.cmet.2019.08.021.

Chung MK, Buck Louis GM, Kannan K, Patel CJ. 2019. Exposome-wide association study of semen quality: systematic discovery of endocrine disrupting chemical biomarkers in fertility require large sample sizes. Environ Int 125:505–514, PMID: 30583854, https://doi.org/10.1016/j.envint.2018.11.037.

Chung MK, Grigoryan H, Iavarone AT, Rappaport SM. 2014. Antibody enrichment and mass spectrometry of albumin-Cys34 adducts. Chem Res Toxicol 27(3):400–407, PMID: 24328277, https://doi.org/10.1021/tx400337k.

Chung MK, Kannan K, Louis GM, Patel CJ. 2018. Toward capturing the exposome: exposure biomarker variability and coexposure patterns in the shared environment. Environ Sci Technol 52(15):8801–8810, PMID: 29972023, https://doi.org/10.1021/acs.est.8b01467.

Chung MK, Patel CJ. 2019. The exposome: an approach toward a comprehensive study of exposures in disease. In: *Encyclopedia of Environmental Health*. Nriagu Jo, ed. 2nd ed. Amsterdam, Netherlands: Elsevier, 770–779.

Chung MK, Regazzoni L, McClean M, Herrick R, Rappaport SM. 2013. A sandwich ELISA for measuring benzo[*a*]pyrene–albumin adducts in human plasma. Anal Biochem 435(2):140–149, PMID: 23333225, https://doi.org/10.1016/j.ab.2012.12.021.

Chung MK, Riby J, Li H, Iavarone AT, Williams ER, Zheng Y, et al. 2010. A sandwich enzyme-linked immunosorbent assay for adducts of polycyclic aromatic hydrocarbons with human serum albumin. Anal Biochem 400(1):123–129, PMID: 20083082, https://doi.org/10.1016/j.ab.2010.01.018.

Czene K, Lichtenstein P, Hemminki K. 2002. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer database. Int J Cancer 99(2):260–266, PMID: 11979442, https://doi.org/10.1002/ijc.10332.

De Vos RCH, Moco S, Lommen A, Keurentjes JJ, Bino RJ, Hall RD. 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc 2(4):778–791, PMID: 17446877, https://doi.org/10.1038/nprot.2007.95.

Dennis KK, Auerbach SS, Balshaw DM, Cui Y, Fallin MD, Smith MT, et al. 2016. The importance of the biological impact of exposure to the concept of the exposome. Environ Health Perspect 124(10):1504–1510, PMID: 27258438, https://doi.org/10.1289/EHP140.

Dunham WH, Mullin M, Gingras A-C. 2012. Affinity-purification coupled to mass spectrometry: basic principles and strategies. Proteomics 12(10):1576–1590, PMID: 22611051, https://doi.org/10.1002/pmic.201100523.

Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, et al. 2011. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc 6(7):1060–1083, PMID: 21720319, https://doi.org/10.1038/nprot.2011.335.

Escher BI, Stapleton HM, Schymanski EL. 2020. Tracking complex mixtures of chemicals in our changing environment. Science 367(6476):388–392, PMID: 31974244, https://doi.org/10.1126/science.aay6636.

Filer D, Patisaul HB, Schug T, Reif D, Thayer K. 2014. Test driving ToxCast: endocrine profiling for 1858 chemicals included in phase II. Curr Opin Pharmacol 19:145–152, PMID: 25460227, https://doi.org/10.1016/j.coph.2014.09.021.

Fischer ES, Böhm K, Lydeard JR, Yang H, Stadler MB, Cavadini S, et al. 2014. Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide. Nature 512(7512):49–53, PMID: 25043012, https://doi.org/10.1038/nature13527.

Florentinus-Mefailoski A, Safi F, Marshall JG. 2014. Enzyme linked immuno mass spectrometric assay (ELIMSA). J Proteomics 96:343–352, PMID: 24316356, https://doi.org/10.1016/j.jprot.2013.11.022.

Go Y-M, Walker DI, Liang Y, Uppal K, Soltow QA, Tran V, et al. 2015a. Reference standardization for mass spectrometry and high-resolution metabolomics applications to exposome research. Toxicol Sci 148(2):531–543, PMID: 26358001, https://doi.org/10.1093/toxsci/kfv198.

Go Y-M, Walker DI, Soltow QA, Uppal K, Wachtman LM, Strobel FH, et al. 2015b. Metabolome-wide association study of phenylalanine in plasma of common marmosets. Amino Acids 47(3):589–601, PMID: 25526869, https://doi.org/10.1007/s00726-014-1893-x.

Grigoryan H, Edmands W, Lu SS, Yano Y, Regazzoni L, Iavarone AT, et al. 2016. Adductomics pipeline for untargeted analysis of modifications to Cys34 of human serum albumin. Anal Chem 88(21):10504–10512, PMID: 27684351, https://doi.org/10.1021/acs.analchem.6b02553.

Ioannidis JPA. 2005. Why most published research findings are false. PLoS Med 2(8):e124, PMID: 16060722, https://doi.org/10.1371/journal.pmed.0020124.

Ioannidis JPA. 2008. Why most discovered true associations are inflated. Epidemiology 19(5):640–648, PMID: 18633328, https://doi.org/10.1097/EDE.0b013e31818131e7.

Ioannidis JPA. 2016. Exposure-wide epidemiology: revisiting Bradford Hill. Stat Med 35(11):1749–1762, PMID: 26646432, https://doi.org/10.1002/sim.6825.

Ioannidis JPA, Loy EY, Poulton R, Chia KS. 2009. Researching genetic versus nongenetic determinants of disease: a comparison and proposed unification. Sci Transl Med 1(7):7ps8, PMID: 20368180, https://doi.org/10.1126/scitranslmed.3000247.

James G, Witten D, Hastie T, Tibshirani R. 2017. *An Introduction to Statistical Learning: With Applications in R*. 1st ed. 2013, Corr. 7th printing 2017 edition. New York, NY: Springer.

Jamin EL, Bonvallot N, Tremblay-Franco M, Cravedi J-P, Chevrier C, Cordier S, et al. 2014. Untargeted profiling of pesticide metabolites by LC–HRMS: an exposomics tool for human exposure evaluation. Anal Bioanal Chem 406(4):1149–1161, PMID: 23892877, https://doi.org/10.1007/s00216-013-7136-2.

Jiang C, Wang X, Li X, Inlora J, Wang T, Liu Q, et al. 2018. Dynamic human environmental exposome revealed by longitudinal personal monitoring. Cell 175(1):277–291.e31, PMID: 30241608, https://doi.org/10.1016/j.cell.2018.08.060.

Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28(1):27–30, PMID: 10592173, https://doi.org/10.1093/nar/28.1.27.

Kempa EE, Hollywood KA, Smith CA, Barran PE. 2019. High throughput screening of complex biological samples with mass spectrometry—from bulk measurements to single cell analysis. Analyst 144(3):872–891, PMID: 30601490, https://doi.org/10.1039/c8an01448e.
Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. 2019. PubChem 2019 update: improved access to chemical data. Nucleic Acids Res 47(D1):D1102–D1109, PMID: 30371825, https://doi.org/10.1093/nar/gky1033.
Klaassen C. 2018. *Casarett & Doull's Toxicology: The Basic Science of Poisons*. 9th ed. New York, NY: McGraw-Hill Education/Medical.
Kortenkamp A. 2014. Low dose mixture effects of endocrine disrupters and their implications for regulatory thresholds in chemical risk assessment. Curr Opin Pharmacol 19:105–111, PMID: 25244397, https://doi.org/10.1016/j.coph.2014.08.006.
Lakhani CM, Tierney BT, Manrai AK, Yang J, Visscher PM, Patel CJ. 2019. Repurposing large health insurance claims data to estimate genetic and environmental contributions in 560 phenotypes. Nat Genet 51(2):327–334, PMID: 30643253, https://doi.org/10.1038/s41588-018-0313-7.
Lanucara F, Holman SW, Gray CJ, Eyers CE. 2014. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nat Chem 6(4):281–294, PMID: 24651194, https://doi.org/10.1038/nchem.1889.
Lazarevic N, Barnett AG, Sly PD, Knibbs LD. 2019. Statistical methodology in studies of prenatal exposure to mixtures of endocrine-disrupting chemicals: a review of existing approaches and new alternatives. Environ Health Perspect 127(2):26001, PMID: 30720337, https://doi.org/10.1289/EHP2207.
Lee H-R, Jeung E-B, Cho M-H, Kim T-H, Leung PCK, Choi K-C. 2013. Molecular mechanism(s) of endocrine-disrupting chemicals and their potent oestrogenicity in diverse cells and tissues that express oestrogen receptors. J Cell Mol Med 17(1):1–11, PMID: 23279634, https://doi.org/10.1111/j.1582-4934.2012.01649.x.
Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859):2224–2260, PMID: 23245609, https://doi.org/10.1016/S0140-6736(12)61766-8.
Lippmann M. 2013. Exposure science in the 21st century: a vision and a strategy. J Expo Sci Environ Epidemiol 23(1):1, PMID: 23242025, https://doi.org/10.1038/jes.2012.109.
Llanos EJ, Leal W, Luu DH, Jost J, Stadler PF, Restrepo G. 2019. Exploration of the chemical space and its three historical regimes. Proc Natl Acad Sci USA 116(26):12660–12665, PMID: 31186353, https://doi.org/10.1073/pnas.1816039116.
Lu Y, Wang Y, Ong C-N, Subramaniam T, Choi HW, Yuan J-M, et al. 2016. Metabolic signatures and risk of type 2 diabetes in a Chinese population: an untargeted metabolomics study using both LC-MS and GC-MS. Diabetologia 59(11):2349–2359, PMID: 27514531, https://doi.org/10.1007/s00125-016-4069-2.
Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, et al. 1995. The nuclear receptor superfamily: the second decade. Cell 83(6):835–839, PMID: 8521507, https://doi.org/10.1016/0092-8674(95)90199-x.
Manrai AK, Cui Y, Bushel PR, Hall M, Karakitsios S, Mattingly CJ, et al. 2017. Informatics and data analytics to support exposome-based discovery for public health. Annu Rev Public Health 38:279–294, PMID: 28068484, https://doi.org/10.1146/annurev-publhealth-082516-012737.
Manrai AK, Ioannidis JPA, Patel CJ. 2019. Signals among signals: Prioritizing nongenetic associations in massive data sets. Am J Epidemiol 188(5):846–850, PMID: 30877292, https://doi.org/10.1093/aje/kwz031.
Mayr LM, Bojanic D. 2009. Novel trends in high-throughput screening. Curr Opin Pharmacol 9(5):580–588, PMID: 19775937, https://doi.org/10.1016/j.coph.2009.08.004.
Medina-Cleghorn D, Bateman LA, Ford B, Heslin A, Fisher KJ, Dalvie ED, et al. 2015. Mapping proteome-wide targets of environmental chemicals using reactivity-based chemoproteomic platforms. Chem Biol 22(10):1394–1405, PMID: 26496688, https://doi.org/10.1016/j.chembiol.2015.09.008.
Meijer J, Lamoree M, Hamers T, Antignac J-P, Hutinet S, Debrauwer L, et al. 2021. An annotation database for chemicals of emerging concern in exposome research. Environ Int 152:106511, PMID: 33773387, https://doi.org/10.1016/j.envint.2021.106511.
Miller GW, Jones DP. 2014. The nature of nurture: refining the definition of the exposome. Toxicol Sci 137(1):1–2, PMID: 24213143, https://doi.org/10.1093/toxsci/kft251.
Morris JH, Knudsen GM, Verschueren E, Johnson JR, Cimermancic P, Greninger AL, et al. 2014. Affinity purification–mass spectrometry and network analysis to understand protein-protein interactions. Nat Protoc 9(11):2539–2554, PMID: 25275790, https://doi.org/10.1038/nprot.2014.164.
Neveu V, Moussy A, Rouaix H, Wedekind R, Pon A, Knox C, et al. 2017. Exposome-Explorer: a manually-curated database on biomarkers of exposure to dietary and environmental factors. Nucleic Acids Res 45(D1):D979–D984, PMID: 27924041, https://doi.org/10.1093/nar/gkw980.
Nisbet ICT, LaGoy PK. 1992. Toxic equivalency factors (TEFs) for polycyclic aromatic hydrocarbons (PAHs). Regul Toxicol Pharmacol 16(3):290–300, PMID: 1293646, https://doi.org/10.1016/0273-2300(92)90009-X.
Osborn MP, Park Y, Parks MB, Burgess LG, Uppal K, Lee K, et al. 2013. Metabolome-wide association study of neovascular age-related macular degeneration. PLoS One 8(8):e72737, PMID: 24015273, https://doi.org/10.1371/journal.pone.0072737.
Patel CJ. 2017. Analytic complexity and challenges in identifying mixtures of exposures associated with phenotypes in the exposome era. Curr Epidemiol Rep 4(1):22–30, PMID: 28251040, https://doi.org/10.1007/s40471-017-0100-5.
Patel CJ, Bhattacharya J, Butte AJ. 2010. An environment-wide association study (EWAS) on type 2 diabetes mellitus. PLoS One 5(5):e10746, PMID: 20505766, https://doi.org/10.1371/journal.pone.0010746.
Patel CJ, Ioannidis JPA. 2014. Studying the elusive environment in large scale. JAMA 311(21):2173–2174, PMID: 24893084, https://doi.org/10.1001/jama.2014.4129.
Patel CJ, Manrai AK. 2015. Development of the exposome correlation globes to map out environment-wide associations. Pac Symp Biocomput 20:231–242, PMID: 25592584, https://doi.org/10.1142/9789814644730_0023.
Peters A, Nawrot TS, Baccarelli AA. 2021. Hallmarks of environmental insults. Cell 184(6):1455–1468, PMID: 33657411, https://doi.org/10.1016/j.cell.2021.01.043.
Polubriaginof FCG, Vanguri R, Quinnies K, Belbin GM, Yahi A, Salmasian H, et al. 2018. Disease heritability inferred from familial relationships reported in medical records. Cell 173(7):1692–1704.e11, PMID: 29779949, https://doi.org/10.1016/j.cell.2018.04.032.
Quinn RA, Nothias L-F, Vining O, Meehan M, Esquenazi E, Dorrestein PC. 2017. Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. Trends Pharmacol Sci 38(2):143–154, PMID: 27842887, https://doi.org/10.1016/j.tips.2016.10.011.
Rappaport SM. 2012. Biomarkers intersect with the exposome. Biomarkers 17(6):483–489, PMID: 22672124, https://doi.org/10.3109/1354750X.2012.691553.
Rappaport SM. 2016. Genetic factors are not the major causes of chronic diseases. PLoS One 11(4):e0154387, PMID: 27105432, https://doi.org/10.1371/journal.pone.0154387.
Rappaport SM. 2018. Redefining environmental exposure for disease etiology. NPJ Syst Biol Appl 4:30, PMID: 30181901, https://doi.org/10.1038/s41540-018-0065-0.
Rappaport SM, Barupal DK, 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, PMID: 24659601, https://doi.org/10.1289/ehp.1308015.
Rappaport SM, Li H, Grigoryan H, Funk WE, Williams ER. 2012. Adductomics: characterizing exposures to reactive electrophiles. Toxicol Lett 213(1):83–90, PMID: 21501670, https://doi.org/10.1016/j.toxlet.2011.04.002.
Rappaport SM, Smith MT. 2010. Environment and disease risks. Science 330(6003):460–461, PMID: 20966241, https://doi.org/10.1126/science.1192603.
Rinschen MM, Ivanisevic J, Giera M, Siuzdak G. 2019. Identification of bioactive metabolites using activity metabolomics. Nat Rev Mol Cell Biol 20(6):353–367, PMID: 30814649, https://doi.org/10.1038/s41580-019-0108-4.
Roca M, Leon N, Pastor A, Yusà V. 2014. Comprehensive analytical strategy for biomonitoring of pesticides in urine by liquid chromatography–orbitrap high resolution mass spectrometry. J Chromatogr A 1374:66–76, PMID: 25499061, https://doi.org/10.1016/j.chroma.2014.11.010.
Roede JR, Uppal K, Park Y, Lee K, Tran V, Walker D, et al. 2013. Serum metabolomics of slow vs. rapid motor progression Parkinson's disease: a pilot study. PLoS One 8(10):e77629, PMID: 24167579, https://doi.org/10.1371/journal.pone.0077629.
Seethala R, Zhang L, eds. 2016. *Handbook of Drug Screening*. 2 ed. New York, NY: CRC Press.
Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM, et al. 2013. Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. Chem Res Toxicol 26(6):878–895, PMID: 23611293, https://doi.org/10.1021/tx400021f.
Smela ME, Hamm ML, Henderson PT, Harris CM, Harris TM, Essigmann JM. 2002. The aflatoxin $B_1$ formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. Proc Natl Acad Sci USA 99(10):6655–6660, PMID: 12011430, https://doi.org/10.1073/pnas.102167699.
Sobus JR, DeWoskin RS, Tan Y-M, Pleil JD, Phillips MB, George BJ, et al. 2015. Uses of NHANES biomarker data for chemical risk assessment: trends, challenges, and opportunities. Environ Health Perspect 123(10):919–927, PMID: 25859901, https://doi.org/10.1289/ehp.1409177.
Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. 2007. Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3(3):211–221, PMID: 24039616, https://doi.org/10.1007/s11306-007-0082-2.
Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. 2019. Benefits and limitations of genome-wide association studies. Nat Rev Genet 20(8):467–484, PMID: 31068683, https://doi.org/10.1038/s41576-019-0127-1.

Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G. 2012. XCMS Online: a web-based platform to process untargeted metabolomic data. Anal Chem 84(11):5035–5039, PMID: 22533540, https://doi.org/10.1021/ac300698c.

Taylor KW, Joubert BR, Braun JM, Dilworth C, Gennings C, Hauser R, et al. 2016. Statistical approaches for assessing health effects of environmental chemical mixtures in epidemiology: lessons from an innovative workshop. Environ Health Perspect 124(12):A227–A229, PMID: 27905274, https://doi.org/10.1289/EHP547.

Tian Z, Zhao H, Peter KT, Gonzalez M, Wetzel J, Wu C, et al. 2021. A ubiquitous tire rubber–derived chemical induces acute mortality in coho salmon. Science 371(6525):185–189, PMID: 33273063, https://doi.org/10.1126/science.abd6951.

Turner MC, Vineis P, Seleiro E, Dijmarescu M, Balshaw D, Bertollini R, et al. 2018. EXPOsOMICS: final policy workshop and stakeholder consultation. BMC Public Health 18(1):260, PMID: 29448939, https://doi.org/10.1186/s12889-018-5160-z.

Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, Yu T, et al. 2013. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. BMC Bioinformatics 14:15, PMID: 23323971, https://doi.org/10.1186/1471-2105-14-15.

Uppal K, Walker DI, Liu K, Li S, Go Y-M, Jones DP. 2016. Computational metabolomics: a framework for the million metabolome. Chem Res Toxicol 29(12):1956–1975, PMID: 27629808, https://doi.org/10.1021/acs.chemrestox.6b00179.

Van den Berg M, Birnbaum L, Bosveld AT, Brunström B, Cook P, Feeley M, et al. 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ Health Perspect 106(12):775–792, PMID: 9831538, https://doi.org/10.1289/ehp.98106775.

Vassilatis DK, Hohmann JG, Zeng H, Li F, Ranchalis JE, Mortrud MT, et al. 2003. The G protein-coupled receptor repertoires of human and mouse. Proc Natl Acad Sci USA 100(8):4903–4908, PMID: 12679517, https://doi.org/10.1073/pnas.0230374100.

Vermeulen R, Schymanski EL, Barabási A-L, Miller GW. 2020. The exposome and health: where chemistry meets biology. Science 367(6476):392–396, PMID: 31974245, https://doi.org/10.1126/science.aay3164.

Vineis P, Chadeau-Hyam M, Gmuender H, Gulliver J, Herceg Z, Kleinjans J, et al. 2017. The exposome in practice: design of the EXPOsOMICS project. Int J Hyg Environ Health 220(2 pt A):142–151, PMID: 27576363, https://doi.org/10.1016/j.ijheh.2016.08.001.

Vrijheid M, Slama R, Robinson O, Chatzi L, Coen M, van den Hazel P, et al. 2014. The Human Early-Life Exposome (HELIX): project rationale and design. Environ Health Perspect 122(6):535–544, PMID: 24610234, https://doi.org/10.1289/ehp.1307204.

Walker DI, Pennell KD, Uppal K, Xia X, Hopke PK, Utell MJ, et al. 2016a. Pilot metabolome-wide association study of benzo(a)pyrene in serum from military personnel. J Occup Environ Med 58(8 suppl 1):S44–S52, PMID: 27501104, https://doi.org/10.1097/JOM.0000000000000772.

Walker DI, Uppal K, Zhang L, Vermeulen R, Smith M, Hu W, et al. 2016b. High-resolution metabolomics of occupational exposure to trichloroethylene. Int J Epidemiol 45(5):1517–1527, PMID: 27707868, https://doi.org/10.1093/ije/dyw218.

Wang A, Gerona RR, Schwartz JM, Lin T, Sirota M, Morello-Frosch R, et al. 2018. A suspect screening method for characterizing multiple chemical exposures among a demographically diverse population of pregnant women in San Francisco. Environ Health Perspect 126(7):077009, PMID: 30044231, https://doi.org/10.1289/EHP2920.

Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, et al. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 472(7341):57–63, PMID: 21475195, https://doi.org/10.1038/nature09922.

Wheelock CE, Rappaport SM. 2020. The role of gene–environment interactions in lung disease: the urgent need for the exposome. Eur Respir J 55(2):1902064, PMID: 32029645, https://doi.org/10.1183/13993003.02064-2019.

Wieser A, Schneider L, Jung J, Schubert S. 2012. MALDI-TOF MS in microbiological diagnostics—identification of microorganisms and beyond (mini review). Appl Microbiol Biotechnol 93(3):965–974, PMID: 22198716, https://doi.org/10.1007/s00253-011-3783-4.

Wild CP. 2005. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomarkers Prev 14(8):1847–1850, PMID: 16103423, https://doi.org/10.1158/1055-9965.EPI-05-0456.

Wild CP. 2012. The exposome: from concept to utility. Int J Epidemiol 41(1):24–32, PMID: 22296988, https://doi.org/10.1093/ije/dyr236.

Wishart D, Arndt D, Pon A, Sajed T, Guo AC, Djoumbou Y, et al. 2015. T3DB: the toxic exposome database. Nucleic Acids Res 43(D1):D928–D934, PMID: 25378312, https://doi.org/10.1093/nar/gku1004.

Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. 2018. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 46(D1):D608–D617, PMID: 29140435, https://doi.org/10.1093/nar/gkx1089.

Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, et al. 2008. DrugBank: a knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res 36(suppl 1):D901–D906, PMID: 18048412, https://doi.org/10.1093/nar/gkm958.

Xia J, Psychogios N, Young N, Wishart DS. 2009. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. Nucleic Acids Res 37(suppl 2):W652–W660, PMID: 19429898, https://doi.org/10.1093/nar/gkp356.

Yang D, Han J, Hall DR, Sun J, Fu J, Kutarna S, et al. 2020. Nontarget screening of per- and polyfluoroalkyl substances binding to human liver fatty acid binding protein. Environ Sci Technol 54(9):5676–5686, PMID: 32249562, https://doi.org/10.1021/acs.est.0c00049.

Yu T, Park Y, Johnson JM, Jones DP. 2009. apLCMS—adaptive processing of high-resolution LC/MS data. Bioinformatics 25(15):1930–1936, PMID: 19414529, https://doi.org/10.1093/bioinformatics/btp291.

Yuan M, Breitkopf SB, Yang X, Asara JM. 2012. A positive/negative ion–switching, targeted mass spectrometry–based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. Nat Protoc 7(5):872–881, PMID: 22498707, https://doi.org/10.1038/nprot.2012.024.

Yuan T-H, Chung M-K, Lin C-Y, Chen S-T, Wu K-Y, Chan C-C. 2016. Metabolic profiling of residents in the vicinity of a petrochemical complex. Sci Total Environ 548–549:260–269, PMID: 26802354, https://doi.org/10.1016/j.scitotenv.2016.01.033.