

## REVIEW

# The exposome and health: Where chemistry meets biology

Roel Vermeulen<sup>1,2\*</sup>, Emma L. Schymanski<sup>3</sup>, Albert-László Barabási<sup>4,5,6</sup>, Gary W. Miller<sup>7\*</sup>

Despite extensive evidence showing that exposure to specific chemicals can lead to disease, current research approaches and regulatory policies fail to address the chemical complexity of our world. To safeguard current and future generations from the increasing number of chemicals polluting our environment, a systematic and agnostic approach is needed. The “exposome” concept strives to capture the diversity and range of exposures to synthetic chemicals, dietary constituents, psychosocial stressors, and physical factors, as well as their corresponding biological responses. Technological advances such as high-resolution mass spectrometry and network science have allowed us to take the first steps toward a comprehensive assessment of the exposome. Given the increased recognition of the dominant role that nongenetic factors play in disease, an effort to characterize the exposome at a scale comparable to that of the human genome is warranted.

A basic tenet of biology is that the phenotype results from a combination of genes and environment. The field of genomics has provided an extraordinary level of genetic knowledge, aided by large-scale, unbiased genome-wide association studies (GWAS). A similar level of analysis, however, is still lacking for the environmental influences on the phenotype. The “exposome” concept was conceived by C. P. Wild in 2005 as a way to represent the environmental, i.e., nongenetic, drivers of health and disease (1). For these external forces to have an effect on health, they must alter our biology, suggesting that a detailed analysis of accessible biological samples at different molecular levels, coupled with information provide snapshots of both the internal (biological perturbations) and external contributors to the exposome. As Rappaport and Smith described in 2010, “toxic effects are mediated through chemicals that alter critical molecules, cells, and physiological processes inside the body ... under this view, exposures are not restricted to chemicals (toxics) entering the body from air, water, or food, for example, but also include chemicals produced by inflammation, oxidative stress, lipid peroxidation, infections, gut flora, and

**“...9 million deaths per year...were attributed to air, water, and soil pollution alone.”**

other natural processes” (2). The exceptional variety and dynamic nature of nongenetic factors (Fig. 1) presents us with an array of sampling and analytical challenges. Fifteen years after the exposome concept was introduced, this review discusses progress in assessing the chemical component of the exposome and its implications on human health.

## From environment to genes

Mapping the human genome revolutionized our ability to explore the genetic origins of disease, but also revealed the limited predictive power of individual genetic variation for many common diseases. For example, genetics contributes to less than half of the risk for heart disease, the leading source of mortality in the United States and many other parts of the world (3). The health impact of environmental risk factors was highlighted by the Global Burden of Disease (GBD) project, which estimated the disease burden of 84 metabolic, environmental, occupational, and behavioral risk factors in 195 countries and territories, and found that these modifiable risks contribute to ~60% of deaths worldwide (4). Using established causal exposure–disease associations, 9 million deaths per year (16% of all deaths worldwide) were attributed to air, water, and soil pollution alone (5). However, the true impact of the environment is likely to be grossly underestimated by these studies, as many of the known chemicals of concern were not considered and less than half of the nongenetic risk burden was explained, suggesting the existence of missing exposome factors (4). These missing factors are analogous to the missing heritability challenge observed in genetic studies. Even with this incomplete inventory, the economic costs of chemical pollution are

considerable, with healthcare and disability-related productivity loss estimated at \$4.6 trillion U.S. dollars per year, representing 6.2% of global economic output (5). Reducing or preventing chemical pollution is a multifaceted problem that involves medical, legal, and regulatory input (see Box 1).

## Measuring chemicals en masse

Several research efforts have pioneered different approaches for the systematic mapping of the exposome, taking advantage of developments in mass spectrometry, sensors, wearables, study design, biostatistics, and bioinformatics (6)—advances that now position us to pursue Dr. Wild’s original vision of the exposome (1). A prime example is how high-resolution mass spectrometry (HRMS) has transformed our ability to measure multitudinous chemical species in a wide range of media, expanding our analytical window beyond targeted analysis of well-known metabolites and priority pollutants (7). HRMS provides the means to simultaneously measure a vast number of exogenous and endogenous compounds, offering a description of the system and its changes in response to exposure to environmental factors (6, 8). As Fig. 2 (top panel) indicates, HRMS is capable of measuring thousands to tens of thousands of chemical features in a single analytical run, although most of these features remain unannotated. Although the systems biology approaches in metabolomics originally focused on detecting endogenous metabolites, HRMS methods can also detect exogenously derived small molecules such as pharmaceuticals, pesticides, plasticizers, flame retardants, preservatives, and microbial metabolites (9). Historically, these exogenous compounds were often viewed as noise and artifact but in reality they carry direct evidence of the complex environments to which living organisms are exposed.

Data resources relevant for HRMS-based exposomics range from specialized lists [e.g., (10)] to medium-sized databases containing tens to hundreds of thousands of chemicals, through to huge resources such as PubChem (11), which has 96 million entries (see Fig. 2). Of the >140,000 chemicals produced and used heavily since the 1950s, only ~5000 are estimated to be dispersed in the environment widely enough to pose a global threat to the human population, although many thousands more would be expected to affect individuals, local communities, or specific occupational settings (5). Specialized lists compiled by, for example, the U.S. Environmental Protection Agency (EPA) (12) and environmental communities such as the NORMAN Network (13) often contain additional information (e.g., exposure data and product information) to help annotate chemicals of interest in the study context. Medium-sized databases such as Human Metabolome Database (HMDB) (14)

<sup>1</sup>Institute for Risk Assessment Sciences, Division of Environmental Epidemiology, Utrecht University, Utrecht, the Netherlands.

<sup>2</sup>Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands.

<sup>3</sup>Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Belvaux, Luxembourg.

<sup>4</sup>Network Science Institute, Northeastern University, Boston, MA, USA.

<sup>5</sup>Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA.

<sup>6</sup>Department of Network and Data Science, Central European University, Budapest, Hungary.

<sup>7</sup>Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, USA.

\*Corresponding author. Email: R.C.H.Vermeulen@uu.nl (R.V.); gm2815@cumc.columbia.edu (G.W.M.)

are commonly used in approaches involving metabolic network analysis, offering typically one to a few possible chemicals per feature of exact mass detected by HRMS. Databases that contain spectral information (i.e., structural “fingerprints”) can be used to increase the confidence of exact mass matching when experimental fragmentation information is available (15, 16). Comprehensive chemical resources such as PubChem are so large that they often offer several thousand possible chemical candidates per exact mass. Despite the exceptional size of the chemical space, the knowledge and computational tools required to interrogate these data are increasingly available (15, 17). For instance, incorporation of literature and patent information with *in silico* methods has greatly improved annotation rates (from <22% to >70%) for >1200 chemicals in HRMS experiments using PubChem (17).

Chemicals are not static entities; they react in our bodies and the environment to form metabolites or transformation products. Computational tools exist to predict such metabolic and environmental transformations (15, 18) but often produce many false-positive and false-negative candidates. Merely predicting first-order reactions of PubChem chemicals would result in billions of possibilities (Fig. 2, second row from bottom). As a result, few studies so far have been able to successfully capitalize on this information in high-throughput identification efforts. The dispersed nature of the essential chemical, metabolite, and spectral information across a wide range of resources with various formats and forms of accessibility (fully open, academic use only, commercial, etc.) is a major impediment to progress in the field.

### Integrating chemical knowledge

The interconnected nature of the available chemical information indicates the need for an interdisciplinary and integrative approach to further define the exposome and the associated data science challenges. Literature mining of PubMed and mapping to discrete chemicals can be used to compile and synthesize the chemical information in the scientific literature (10, 19). The expansion and automation of literature mining for more accurate chemical candidate retrieval during high-throughput identification, e.g., with MetFrag (17) or other *in silico* approaches (15), will be crucial for faster, more efficient annotation of the complex and highly varying datasets that characterize studies of chemical exposures and health.

Many of the chemicals of interest in exposomics come from the same or related sources (e.g., industrial processes, consumer goods, diet), meaning that such exposures exhibit a population structure (i.e., complex correlations and dynamic patterns) akin to observed cor-

relations in complex biological systems. Thus, the reduction of dimensional complexity will be possible by grouping correlated exposures. Indeed, several reports have shown correlation patterns between different chemicals and chemical families within populations (20, 21). These relationships between chemicals can be presented as networks of chemicals (i.e., exposure enrichment pathways) that reveal communities of exposures (20, 21), which in turn can be used to explore the impact that they have on the biological system (see the following section).

Much of our current knowledge about the health effects of chemicals comes from epidemiological and toxicological studies in which a few pollutants are analyzed in relation to a specific phenotype, representing a hypothesis-

driven path toward understanding exposure-disease relationships. However, our exposures are not a simple sum of a handful of chemicals. To overcome the limitations of traditional epidemiological studies, environment-wide association studies (EWAS) have been proposed for identifying new environmental factors in disease and disease-related phenotypes at scale. EWAS was inspired by the analytical procedures developed in GWAS (22) in which a panel of “exposures,” analogous to genotype variants, is studied in relation to a phenotype of interest. For example, using the National Health and Nutrition Examination Survey dataset, an EWAS study explored the associations of 543 environmental attributes with type 2 diabetes, identifying five statistically

### Ecosystems

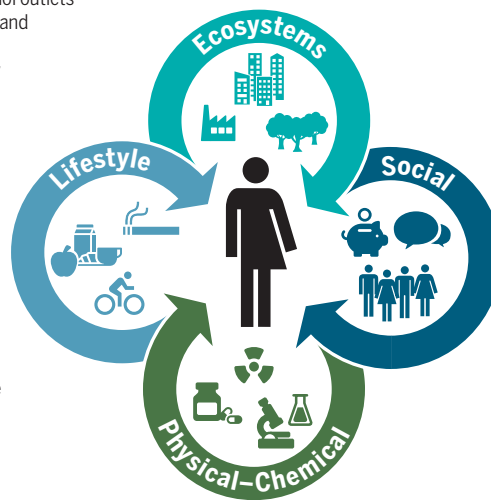
Food outlets, alcohol outlets  
Built environment and urban land uses  
Population density  
Walkability  
Green/blue space

### Lifestyle

Physical activity  
Sleep behavior  
Diet  
Drug use  
Smoking  
Alcohol use

### Social

Household income  
Inequality  
Social capital  
Social networks  
Cultural norms  
Cultural capital  
Psychological and mental stress



### Physical-Chemical

Temperature/humidity  
Electromagnetic fields  
Ambient light  
Odor and noise  
Point, line sources, e.g. factories, ports  
Outdoor and indoor air pollution  
Agricultural activities, livestock  
Pollen/mold/fungus  
Pesticides  
Fragrance products  
Flame retardants (PBDEs)  
Persistent organic pollutants  
Plastic and plasticizers  
Food contaminants  
Soil contaminants  
Drinking water contamination  
Groundwater contamination  
Surface water contamination  
Occupational exposures

**Fig. 1. The exposome concept.** The exposome is an integrated function of exposure on our body, including what we eat and do, our experiences, and where we live and work. The chemical exposome is an important and integral part of the exposome concept. Examples of external stressors are adapted from (39). These stressors are reflected in internal biological perturbations (Fig. 3); therefore, exposures are not restricted to chemicals (toxicants) entering the body, but also include chemicals produced by biological and other natural processes.

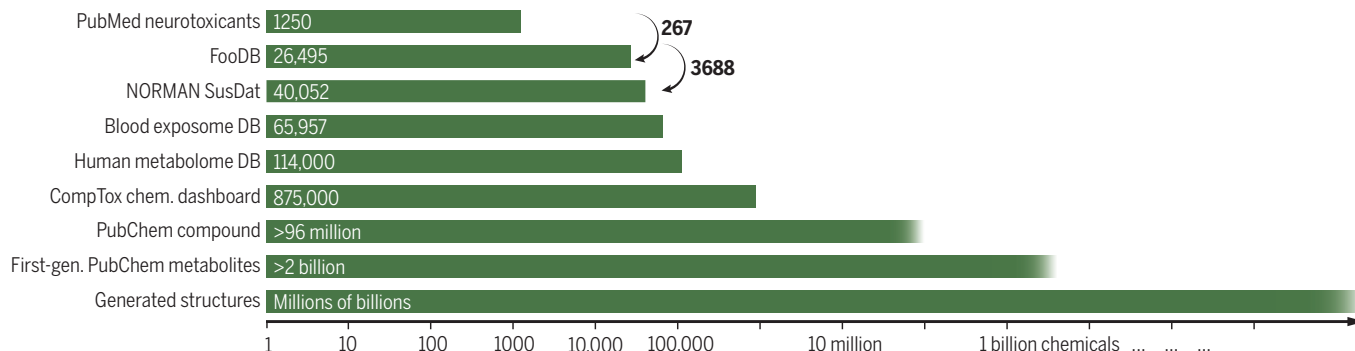
### Box 1. The exposome and regulation.

Many of the influential regulatory bodies in Europe and North America have been expanding their computational and high-throughput approaches to address the increasing number of chemicals to which humans are exposed, but there are still major challenges regarding prioritization. Networks such as NORMAN (13), which bridge scientists, regulators, and practitioners, are becoming increasingly valuable avenues of knowledge exchange. Large-scale exposome studies provide a systematic approach to prioritization, allowing regulatory bodies to focus on those chemicals that have the largest adverse effects on health. If systematic analysis reveals major adverse effects on human health from exposure to currently approved or potential replacement chemicals, then those compounds should be removed from the marketplace. Although thousands of compounds are classified as “generally recognized as safe,” they have never been subjected to the scientifically rigorous testing systems currently in place. A data-driven exposome approach ignores historical decision-making and can help to evaluate the effects of classes of chemicals on specific biological pathways known to be perturbed, which will help in the design of new compounds with minimal impact on human health and the environment.

### Typical HRMS sample



### Selected exposomics, chemical data sources



**Fig. 2. Chemical complexity of HRMS and the exposome.** Top: Known versus unknown features in a typical HRMS measurement [data from (7)]. Bottom: Selected data sources relevant to the chemical exposome (10–14, 19). Arrows show the overlap of potential neurotoxics in FooDB (<http://foodb.ca/>) and FooDB components in NORMAN SusDat ([www.norman-network.com/nds/susdat/](http://www.norman-network.com/nds/susdat/)) (prioritized chemicals of environmental interest).

significant associations (including persistent organic pollutants and pesticides) validated across independent cohorts (22). However, by focusing on a predetermined list of chemicals, these initial EWAS studies likely suffer from the same limitations of candidate gene searches. Further, current EWAS approaches do not test for interactions and/or combinations of factors (mixtures). Recent efforts have been undertaken to develop statistical methods to screen for interactions and test the effects of mixtures or to apply frameworks such as aggregated exposure and adverse outcome pathways to study combinatorial effects (9).

As systematic exposomics moves forward to elucidate the impact of the constellation of chemical exposures on our health, increasingly rich and high-dimensional data must be captured (Fig. 3). In addition, defining the appropriate frameworks for establishing controls, as well as background and negative responses, is essential for enabling causal inference. To aid inference, more insights into the boundaries of what are “normal” responses are required and necessitate definitions of a reference exposome.

#### Network science to address exposome complexity

The challenge in understanding the role of the exposome on our health lies not only in the large number of chemical exposures in our daily lives, but also in the complex ways that they interact with cells. A reductionist approach might isolate the role of a single variable, but it will inadequately capture the complexity of

the exposome. Network science (23), which has well-developed applications in medicine and systems biology (24), offers a platform with which to achieve an understanding of the impact of multiple exposures. Each chemical will exert its effect through interactions with various cellular components supplying or perturbing cellular networks. To capture the diversity in these interactions, we must first catalog the sum of all physical interactions as a multilayer network (25) consisting of several distinct biological layers (Fig. 3). Although each of these networks will rely on different biological mechanisms, they are not independent; for example, protein production is governed by the regulatory network, and the catalysis of the metabolic reactions is in turn governed by the enzymes and protein complexes of the regulatory network (26).

To fully understand the role of the exposome, we must similarly develop a multilayer network-based framework capable of unveiling the role of chemicals, their combinations, and biological perturbations on our health. However, there are several data and methodological challenges. The first challenge is the paucity of systematic data on the various dimensions of exposure, from bioavailability to protein-binding information of the hundreds of thousands of exposome molecules. The U.S. National Toxicology Program, the EPA, and the European Molecular Biology Laboratory (EMBL) are developing platforms to generate, collate, and organize data on chemical-biological interactions, but there is a need for high-throughput approaches that

offer greater coverage (12, 27, 28). The second challenge in developing a framework is that the current statistical toolset assumes that we are faced with a collection of random variables that are independent, identically distributed, and measured with equal precision. In a network environment, these assumptions are inherently false, as interactions couple the probability distribution of most network-based variables. Furthermore, most of the chemicals we are exposed to represent communities of exposures, so the effect of a chemical is rarely observed in isolation. Therefore, identifying meaningful associations from high-dimensional exposomic data poses major statistical and computational challenges that need to be addressed in parallel with experimental developments. The third challenge is that, beyond cataloging interactions, we must also understand the dynamics of the biochemical pathways (29) through which different elements of the exposome affect our health. Indeed, the human interactome, representing the sum of all physical interactions within a cell (Fig. 3), is often depicted as a static graph but is in reality a temporal network (30) with nodes and links that disappear and reemerge depending on factors ranging from the cell cycle to variability in environmental exposures across the life course. Modeling the fully temporal nature of these networks remains a challenge, as the kinetic constants underlying metabolic processes are not known and we currently lack systematic tools with which to identify them (31).

#### Informative exposome study designs

A systematic and unbiased assessment of the exposome that does not focus on a selected set of readily measured or priority chemicals requires access to biological samples that

**“Network science...offers a platform...to [understand]... the impact of multiple exposures.”**

reflect exposures, biological effects, and, preferably, the health phenotype of interest. This is challenging because it will be rare that the variability of exposures (E) aligns perfectly with the kinetics of the biological effects (B) or the etiological time window of the health phenotype, including developmental and trans-generational effects (P). Optimizing each step (E-B and B-P) in separate studies, however, has the disadvantage that overlapping patterns in each step restrict us from unveiling the true association between exposure and the health phenotype (E-P). The meet-in-the-middle (MITM) design attempts to address this challenge (32). In MITM, exposures can be assessed in individuals using HRMS or upstream estimates of external factors (Fig. 1) and are compared with downstream biological changes in persons who develop a specific health phenotype and those who do not.

The MITM approach using HRMS data has successfully identified single and combinatorial effects of chemicals (33–36). For example, the HELIX study explored the early-life exposome of population-based birth cohorts and identified several environmental chemicals that were associated with lung function in children (35). The EXPOsOMICS study showed how air pollution alters biological pathways, particularly linoleate metabolism, which predicted the occurrence of adult-onset asthma and cardiovascular disease (36).

### Scaling up

By pooling studies, sample sizes for GWAS have increased from a few thousand to tens to hundreds of thousands of individuals over the past decade (37). However, enrollment in studies of nongenetic environmental exposures remains relatively low. The large-scale genomic consortia efforts allowed GWAS to detect many common genetic traits related to health phenotypes and, although the combined effects of the identified traits are still modest, they provide insights into the underlying biological pathways of disease. It is estimated that sample sizes of 500,000 to 2,000,000 are needed to explain a substantial portion of the projected genomic heritability of common chronic diseases (38). For the multitude of factors within the exposome, most of which likely exert small effects, similar or even greater sample sizes would be required for future environmental studies and EWAS (22). Scaling exposome research to these numbers will require a joint effort across multiple cohort consortia and research programs. Recently funded programs to work toward a human exposome project are a first step toward reaching tens of thousands of people with detailed environmental and biological analysis of exposures. Although these numbers are large enough to identify the most prevalent and strongest chemical risk factors, progressive increments in sample size will be

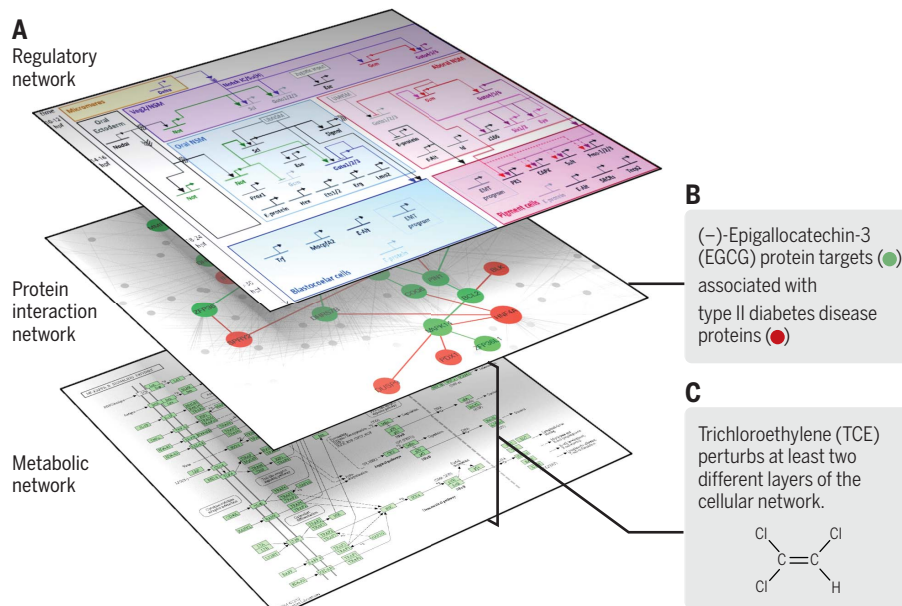
needed for a systematic understanding of the impact of combinatorial exposome factors on specific and rare phenotypes. The systematic identification of the impact of nongenetic factors and chemical exposures would enable the establishment of an exposome risk score (ERS) akin to the polygenic risk score (PRS) (see Box 2).

### Next steps for the exposome

The rate, volume, and variety of chemicals being introduced into our environment continue to expand. The importance of these chemical exposures on human health is exemplified by the large proportion of disease caused by as yet unknown exposome factors (3). Indeed, the nongenetic component exceeds known and missing heritability. We therefore need innovative ways to study these factors and translate our findings into policy. Currently, many of the regulatory agencies are expanding their computational and high-throughput approaches

to account for the ever-increasing number of chemicals, but there are still major challenges regarding prioritization and new approaches are urgently needed (see Box 1). Open science efforts such as Global Natural Product Social Molecular Networking (GNPS), which allow users to archive huge amounts of raw data and in return offers computational mass spectrometry workflows coupled with open mass spectral libraries and continuous updates of new identifications, are beginning to be leveraged for large-scale studies (20). However, as discussed above, we must address several challenges to exploit the full potential of exposome research as it relates to improving our understanding of exposure to complex chemical mixtures. To address these challenges, we must: (i) improve our technology to screen for exogenous chemicals and their biological consequences at higher-throughput rates and lower costs; (ii) continue to develop

### The cell as a multilayer network



**Fig. 3. Impact of the exposome on subcellular networks.** (A) Network medicine views the cell as a multilayer network with three principal, interdependent layers: (i) a regulatory network capturing all interactions affecting RNA and protein expression, (ii) a protein interaction network that captures all binding interactions responsible for the formation of protein complexes and signaling, and (iii) a metabolic network representing all metabolic reactions, including those derived from the microbiome, a network of interacting bacteria linked through the exchange of metabolites. Exposome-related factors can affect each layer of this multilayer network. (B) For example, the polyphenol epigallocatechin gallate (EGCG), a biochemical compound in green tea with potential therapeutic effects on type 2 diabetes mellitus (T2D), binds to at least 52 proteins (40). Network-based metrics reveal a proximity between these targets and 83 proteins associated with T2D, suggesting multiple mechanistic pathways to potentially account for the relationship between green tea consumption and reduced risk of T2D. (C) As another example, trichloroethylene (TCE) is a volatile organic compound that was widely used in industrial settings and is now a widespread environmental contaminant present in drinking water, indoor environments, ambient air, groundwater, and soil. Multiple lines of evidence support a link between TCE exposure and kidney cancer and possibly non-Hodgkin's lymphoma (33). TCE perturbs at least two different layers of the cellular network: It covalently binds to proteins from the protein interaction network, altering their function, and affects the cellular metabolic network, eventually leading to adenosine triphosphate (ATP) depletion. Network-based tools could be used to explore the mechanistic role of many other exposome chemicals on our health and to build experimentally testable hypotheses.

the current chemical and spectral data resources needed to identify these signals in samples; (iii) increase the scale and scope of studies to a level that provides the necessary statistical power to precisely characterize the effects of the chemicals and their combinations; (iv) further develop and support cheminformatic and bioinformatic tools, including network theory and network medicine, to elucidate the constellation of the chemical environment and its biological consequences; and (v) ensure adequate protection for the generation of false-positive results by insisting on replication in independent studies and the use of methods to establish causation, such as Mendelian randomization, within-sibling comparisons, and exposure-negative and outcome-negative controls.

### Box 2. Toward an ERS.

There has been substantial progress in the identification of genetic risk factors for chronic diseases. Analysis of high-risk mutations and estimation of PRS for these diseases are now becoming routine and can be included when developing individual-based (i.e., precision) prevention and treatment strategies. Similarly, the establishment of ERS would help to summarize relevant nongenetic risk factors, enabling the identification of hotspots of concern where multiple environmental factors come together, and would aid in the prioritization of risk factors on the basis of their population and individual impact. For example, an ERS could provide data on exposure to chemical toxicants that are based on the biological processes or organ systems that are most vulnerable and couple them with indices of associated biological injury or response. Such ERS, in contrast to PRS, would be time varying and dynamic through age-related exposures and susceptibilities.

### Conclusion

A concerted and systematic effort to profile the nongenetic factors associated with disease and health outcomes is urgently needed because we lack important insights that might assist us in curtailing the ever-growing burden of chronic disease on society. Emerging exposome research frameworks are poised to enable the systematic analysis of nongenetic factors involved in disease. Technology has enabled the first generation of studies to evolve into the comprehensive study of combinatorial chemical exposures. Future efforts must ensure that analytical approaches and study designs are rigorous and validated. A coordinated and international effort to characterize the exposome, akin to the Human Genome Project, would provide rigorous data to allow exposome-based EWAS to be conducted at the scale of GWAS. By taking advantage of the nontargeted nature of HRMS, EWAS provide a true complement to GWAS. Consolidating knowledge garnered from GWAS and EWAS would allow us to map the gene and environment interface, which is where nature meets nurture and chemistry meets biology.

### REFERENCES AND NOTES

- C. P. Wild, *Cancer Epidemiol. Biomarkers Prev.* **14**, 1847–1850 (2005).
- S. M. Rappaport, M. T. Smith, *Science* **330**, 460–461 (2010).
- J. P. Ioannidis, E. Y. Loy, R. Poulton, K. S. Chia, *Sci. Transl. Med.* **1**, 7ps8 (2009).
- GBD 2016 Risk Factors Collaborators, *Lancet* **390**, 1345–1422 (2017).
- P. J. Landrigan *et al.*, *Lancet* **391**, 462–512 (2018).
- M. M. Niedzwiecki *et al.*, *Annu. Rev. Pharmacol. Toxicol.* **59**, 107–127 (2019).
- E. L. Schymanski *et al.*, *Environ. Sci. Technol.* **48**, 1811–1818 (2014).
- D. C. Sévin, A. Kuehne, N. Zamboni, U. Sauer, *Curr. Opin. Biotechnol.* **34**, 1–8 (2015).
- B. I. Escher *et al.*, *Environ. Int.* **99**, 97–106 (2017).
- E. L. Schymanski *et al.*, *Environ. Sci. Process. Impacts* **21**, 1426–1445 (2019).
- S. Kim *et al.*, *Nucleic Acids Res.* **47** (D1), D1102–D1109 (2019).
- A. J. Williams *et al.*, *J. Cheminform.* **9**, 61 (2017).
- V. Dulio *et al.*, *Environ. Sci. Eur.* **30**, 5 (2018).
- D. S. Wishart *et al.*, *Nucleic Acids Res.* **46** (D1), D608–D617 (2018).
- I. Blaženović, T. Kind, J. Ji, O. Fiehn, *Metabolites* **8**, 31 (2018).
- E. L. Schymanski *et al.*, *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
- C. Ruttkies, E. L. Schymanski, S. Wolf, J. Hollender, S. Neumann, *J. Cheminform.* **8**, 3 (2016).
- Y. Djoumbou-Feunang *et al.*, *J. Cheminform.* **11**, 2 (2019).
- D. K. Barupal, O. Fiehn, *Environ. Health Perspect.* **127**, 097008 (2019).
- J. M. Gauglitz *et al.*, *Food Chem.* **302**, 125290 (2020).
- S. Li *et al.*, *Reprod. Toxicol.* S0890-6238(18)30603-8 (2019).
- C. J. Patel, J. Bhattacharya, A. J. Butte, *PLoS ONE* **5**, e10746 (2010).
- A.-L. Barabási, *Network Science* (Cambridge Univ. Press, 2016).
- J. Loscalzo, A.-L. Barabási, E. K. Silverman, *Network Medicine: Complex Systems in Human Disease and Therapeutics*. (Harvard Univ. Press, 2017).
- G. Bianconi, *Multilayer Networks: Structure and Function* (Oxford Univ. Press, ed. 1, 2018).
- A.-L. Barabási, Z. N. Oltvai, *Nat. Rev. Genet.* **5**, 101–113 (2004).
- D. Mendez *et al.*, *Nucleic Acids Res.* **47** (D1), D930–D940 (2019).
- R. R. Tice, C. P. Austin, R. J. Kavlock, J. R. Bucher, *Environ. Health Perspect.* **121**, 756–765 (2013).
- A. Barrat, M. Barthélemy, A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge Univ. Press, 2008).
- P. Holme, J. Saramäki, "Temporal networks as a modeling framework," in *Understanding Complex Systems* (Springer, 2013), pp 1–14.
- M. Santolini, A.-L. Barabási, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6375–E6383 (2018).
- M. Chadeau-Hyam *et al.*, *Environ. Mol. Mutagen.* **54**, 542–557 (2013).
- D. I. Walker *et al.*, *Int. J. Epidemiol.* **45**, 1517–1527 (2016).
- B. Warth *et al.*, *Anal. Chem.* **89**, 11505–11513 (2017).
- L. Agier *et al.*, *Lancet Planet. Health* **3**, e81–e92 (2019).
- A. Jeong *et al.*, *Environ. Int.* **119**, 334–345 (2018).
- T. Beck, R. K. Hastings, S. Gollapudi, R. C. Free, A. J. Brookes, *Eur. J. Hum. Genet.* **22**, 949–952 (2014).
- Y. Zhang, G. Qi, J. H. Park, N. Chatterjee, *Nat. Genet.* **50**, 1318–1326 (2018).
- M. C. Turner *et al.*, *Annu. Rev. Public Health* **38**, 215–239 (2017).
- H. Iso, C. Date, K. Wakai, M. Fukui, A. Tamakoshi; JACC Study Group, *Ann. Intern. Med.* **144**, 554–562 (2006).

### ACKNOWLEDGMENTS

We thank the following colleagues for critical review of this manuscript: R. Balling, M. Chadeau-Hyam, G. Downward, L. P. Fried, D. P. Jones, V. Kalia, V. Lenters, G. Menichetti, R. Singh, I. Valle, B. van de Water, and J. Vlaanderen. **Funding:** R.V. is supported by an EU H2020-EXPANSE grant, the NWO Gravitation Program, and intramural funding from Utrecht University. E.L.S. is supported by the Luxembourg National Research Fund (FNR grant no. A18/BM/12341006). G.W.M. is supported by the National Institutes of Health (NIH grant nos. U2ESCO30163 and RC2 DK118619). A.-L.B. is supported by the NIH (grant no. P01HL132825) and the American Heart Association (grant no. 151708). **Competing interests:** A.-L.B. is founder of Nomix, Foodome, and Scipher Medicine, companies that explore the role of networks in health. The other authors declare no competing interests. **Data and materials availability:** All data are available in the main text.

10.1126/science.aay3164

## The exposome and health: Where chemistry meets biology

Roel Vermeulen, Emma L. Schymanski, Albert-László Barabási and Gary W. Miller

*Science* **367** (6476), 392-396.  
DOI: 10.1126/science.aay3164

### ARTICLE TOOLS

<http://science.sciencemag.org/content/367/6476/392>

### RELATED CONTENT

<http://science.sciencemag.org/content/sci/367/6476/378.full>  
<http://science.sciencemag.org/content/sci/367/6476/380.full>  
<http://science.sciencemag.org/content/sci/367/6476/384.full>  
<http://science.sciencemag.org/content/sci/367/6476/388.full>  
<http://science.sciencemag.org/content/sci/367/6476/397.full>  
<http://science.sciencemag.org/content/sci/367/6476/360.full>  
<http://science.sciencemag.org/content/sci/367/6476/369.full>

### REFERENCES

This article cites 35 articles, 4 of which you can access for free  
<http://science.sciencemag.org/content/367/6476/392#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

---

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2020, American Association for the Advancement of Science