Feature article:

The evolution of environmental exposure science: Using breath-borne biomarkers for "discovery" of the human exposome

Joachim D. Pleil[1]* and Matthew A. Stiegel[2]

[1]National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA
[2]Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, 27599, USA

According to recent research, 70-90% of long-term latency and chronic human disease incidence is attributable to environmental (human exposome) factors through the gene - environment interaction. Environmental exposure science is now embarking on a new "discovery" path for decoding the human exposome using biomarkers in breath and other biological media.

*corresponding author: pleil.joachim@epa.gov

Environmental exposure science and the human exposome

Exhaled breath contains much more than gases; the aerosols carry along proteins, cell fragments, bacteria, DNA, viruses and other large molecules.

Recent developments in environmental research have focused on the cumulative effects of the total environment on human metabolism and health state. The fundamental discipline, Exposure Science, is now being established as an academic specialty distinct from toxicology, epidemiology, and biology.[1] Exposure Science has been defined as studying *"…human contact with chemical, physical, or biological agents occurring in their environments, (to) advance knowledge of the mechanisms and dynamics of events either causing or preventing adverse health outcomes…"* and is considered the "*…bedrock for protection of public health*".[2-4] In short, Exposure Science encompasses the study of the external environment's contributions to the human system as well as the chemicals already present within human biological media; these are collectively referred to as the human exposome.[5,6] The study of the human exposome is now gaining momentum; for example the European Helix Project is focusing on the "early-life exposome", the U.S. National Institutes of Environmental Health Sciences (NIEHS) and U.S. Environmental Protection Agency (EPA) have jointly released a report entitled "Exposure Science in the 21st Century: A vision and strategy" for exploring the links between environment and disease, the NIEHS has established a Center for Exposure Biology at University of California, Berkeley and the first U.S. Human Exposome Center has recently been funded at Emory University in Atlanta, Georgia.[7-10]

The general concept of the exposome considers all of the chemicals in the body representing environmental exposures received by an individual over a lifetime, as well as chemical metabolites, metabolic reaction products, byproducts of the microbiome, and systemic response molecules. In short, the exposome represents everything that is "not the genome" in human systems biology.[5,11-13] However, in contrast to the genome, the exposome fluctuates throughout the human lifetime as it is constantly influenced by changes in the environment, human activities, and the individual's current metabolism and health state at the molecular, cellular and organ level of organization representing the overall concept of human systems biology.[14-16] So, why do we care about such a complicated and moving target?

We care because current research is showing that incidence of cancers, as well as autoimmune, cardiovascular, pulmonary, and other chronic diseases are not just bad luck,

but that 70-90% of such human disease is attributable to environmental (human exposome) factors through the gene-environment interaction and the epigenome.[11,17,18] This has appeared as a bit of a surprise because the sequencing of the human genome, announced in 2000, and the subsequent associations of function with specific genes were thought to be the final explanation for human disease incidence.[19] Instead, the genome results and expected unambiguous linkages to disease have been elusive.[20-22]

To understand the underlying issue, consider familial relatives such as a grandfather and grandson pair; on average, they have a probability of sharing 25% of the same genetic traits. However, they have very little likelihood of sharing the same exposome because the grandfather has two generations of lifetime exposure head start. Furthermore, only when he and the grandson are collocated as in the picture above, do they share the same short-term exposures, and even then, they are likely to absorb exogenous chemicals at different rates and process them with different efficiencies. In short, there is little expectation that their exposomes are as similar as their genomes over time and so their long term health outcomes are much more likely to be different. In contrast, consider contemporary occupants of a locality or other kind of community that have grown-up and lived for a long time in the local environment. If long-term health outcome is indeed driven by the exposome as suggested by the research, then the long times spent in a common local environment should be much more important than whether or not the people are genetically related.

Relatives like the grandfather and grandson above share genetic traits, but their environmental lifetime experiences result in very different personal exposomes despite sharing the same exposures for a short time on occasion.

Targeted biomarker research

This article is concerned with strategies for teasing out the constituents of the exposome that are significant for deducing the links between the environment, internal dose, and the apparently random outcomes in human health. The traditional approaches have centered on chemicals known to have engendered adverse health effects and studying linkages between environmental levels and internal doses; these are referred to as "knowledge driven investigations". Some examples of such methods are the measurements of "benzene in the air *vs.* benzene in the breath", or "chlorpyrifos pesticide in the food *vs.* trichloropyridinol metabolite in the urine".[23,24] The gist of these contrasts is to estimate how much of a known hazardous environmental contaminant is getting into the human system, being distributed to target organs, and subsequently metabolized, and then trying to use various calculations from animal and *in vitro* toxicological studies to assign health

risk. This is called "targeted" or "bottom-up" research and has served us well in regulating chemicals known to affect public health.[17] The problem is that there are over 1,000 new chemicals introduced into the U.S. environment every year (from consumer products, industrial processes, farming operations, etc.) to join the more than 84,000 chemicals already listed in the Toxic Substances Control Act inventory.[25] Targeted research has only been performed for a few hundred of the presumed most toxic and/or prevalent of these chemicals; there are limited toxicity data of any kind for many of these compounds, and so the toxicological assessment system is becoming overwhelmed.[26-28]

Certainly, performing targeted *in vivo* research on all potentially toxic chemicals is not practical; the risk assessment community is exploring alternate paths especially for prioritizing chemicals for more detailed study.[29] The newer approaches implement *in vitro* (cellular/molecular level) and *in silico* (computational) methods for screening chemicals that promise high throughput results on the order of thousands of chemicals per year.[30,31] However, despite an increase in throughput, these methods still fall into the targeted category wherein *in vivo* experiments concerning systems biology, mode of action (MoA), adverse outcome pathways (AOP), and biochemical reactions are required.[32-34] In short, targeted assessments require prior knowledge which is best developed through discovery without preconceptions.

Discovery biomarker research

Targeted toxicity testing is unquestionably valuable for delineating specific risks from specific compounds based on prior knowledge. The inverse procedure, wherein we try to discover which environmental chemicals, human metabolites, and endogenous biological compounds are related to pre-clinical and ultimately adverse effects, requires a different strategy. The study of "exposure biology" is part of the broader discipline of Exposure Science that encompasses the discovery aspects and has been described as follows:

> *"Epidemiologists wait for people to die or get sick before they can study them. Exposure biology connects exposures to hazardous chemicals with early effects of these exposures inside the body." (Stephen Rappaport, Director, Center for Exposure Biology, University of California, Berkeley).*

The immediacy of the pre-clinical effect reduces a great deal of uncertainty in making the connections to long-term latency disease occurrence. This prompts us to discover what biomarker patterns are prospective signs of disease, and/or retrospective signs of exposure. In contrast to *in vitro* and *in silico* strategies, discovery generally requires samples from intact biological systems (human, mouse, rat, etc.). We are also working on discovery analyses using cultured bacteria representing the human microbiome which is technically *in vitro*, however, this still represents intact, albeit very small, organisms. We note that discovery methods do not necessarily preempt high throughput analysis for different chemicals because we are looking at large groups of biological compounds simultaneously.

So, how does exposome discovery actually work? In the figure below, we show a conceptual diagram contrasting the classical concept of targeted and discovery approaches to Exposure Science. In the targeted strategy, we start with a knowledge base of specific hazardous exogenous chemicals, and then look for them in environmental media such as food, air, and water to determine what the potential of exposures could be. In the discovery approach, we start without preconceived notions about chemicals and just begin to analyze human biofluids (e.g. blood, breath, urine) for as many chemicals as we can. Unlike discovery within the genetics community where a single analytical platform such as a DNA microarray, or "Affy chip" (Affymetrix, Santa Clara CA) exists that can give results from tens of thousands of probes for all specific sequences of interest simultaneously[31], there is no single chemical analytical system (to date) that can analyze every kind of chemical at once.

Top-down "discovery" analysis of all chemicals in blood, breath and urine

Bottom-up "targeted" analysis of compounds in environmental

The diagram above (adapted with permission from Prof. Steven Rappaport from University of California, Berkeley) shows the classical concepts of bottom-up, targeted analysis (left triangle) in contrast to top-down discovery analysis (right triangle) for assessing the effects on the human community or sub-population (center). Targeted analysis seeks out specific compounds of exposure in the environment from multiple sources (food, air, water, dust, etc.) whereas discovery analysis starts with human biological media (blood, breath, urine, etc.) to look for all possible compounds and then narrow them down to those with potential exposure to health linkages.

This classical concept described in the figure above is not necessarily practical or applicable in all cases. Most discovery experiments are targeted to some extent in that different compound types require different analytical approaches. Most targeted experiments fall partially into the discovery arena because when we target a group of compounds, we can subsequently "discover" that some turn out to be inconsequential for the particular hypothesis. `Also, discovery and targeted analyses can be done within the same sample type; that is we can discover many compounds in exhaled breath, and then subsequently target an important subset within breath samples without invoking the environment at all. Conversely, we can perform discovery experiments using

environmental media samples to deduce what may be worthwhile targeting for assessing potential environmental toxicity in air, water or food.

Exposome discovery requires a wide array of analytical techniques and sample processing methods to accommodate the different types and categories of bio-chemicals in the human body. We could not expect to measure small reactive compounds, non-polar exogenous chemicals, endogenous polar compounds and proteins all with the same methods. The best we can do is to pick our battles in a stratified fashion; gas chromatography – mass spectrometry (GC-MS) for volatiles, liquid chromatography mass-spectrometry (LC-MS) for larger polar molecules, immunochemistry for proteins, etc. with appropriate sample collection, storage, and preparation to accommodate the medium (blood, breath, or urine). Similarly, we have the same practical issues for environmental samples, and as such, the strict division between discovery and targeted analysis is always blurred to some extent.

Environment wide association study (EWAS) and case-control discovery

The recent article by Patel et al. was heralded as a true exposome discovery experiment conducted using a genome-wide association study (GWAS) approach.[32] Dubbed "EWAS" wherein the word "environment" replaces "genome", this style of research uses chemical and environmental biomarkers as the independent variables instead of the loci from sequenced genetic samples.[33] Patel et al. used 266 unique environmental factors from the National Health and Nutrition Examination Survey (NHANES) as potential associations with clinical status of diabetes mellitus as indicated by fasting blood sugar levels >126 mg/dl. They found significant positive associations with the pesticide derivative heptachlor epoxide, the vitamin γ-tocopherol, and polychlorinated bi-phenyls, and protective factors associated with β-carotenes. As discussed above, this prototype exposome discovery experiment can also be considered to be targeted to some extent in that the 266 factors used were initially chosen by CDC to accommodate their sampling and analytical strategies and so do not represent all compounds available in the biological media. Regardless of these finer distinctions, such exposome discovery studies can be used to develop hypotheses as to the causes of different diseases, and perhaps even provide pre-clinical evidence as to early stage susceptibility.

More prospective approaches have been performed on a smaller scale over the years under the umbrella of "case-control" studies. Herein, specific sets of compounds are chosen *a priori* based on available instrumentation and access to biological media, and then differences in their concentration patterns between a group of individuals with a diagnosed illness and a group of matched control subjects are investigated. For example, there have been a number of exhaled breath studies for trying to discern pre-clinical markers for lung cancer. This is of particular importance as early detection radically improves survival rate but symptoms are often mild or ambiguous during early stages of the cancer and so no specific imaging or biopsies are indicated.[34-37] For lung cancer especially, it seems that there would be some exhaled gas-phase chemicals that would give a clear early warning, but discovering what these chemicals could be is as yet difficult.

In both cases, the approaches have a discovery nature in that we do not make assumptions, but rather let the chips (chemicals) fall where they may. The problem with such biomarker strategies is generalizability and repeatability. Often, an algorithm based on a pattern of biomarkers is able to separate the cases from the controls in the current experiment only to fail when the experiment is repeated with different subjects. This is often the case when there are confounding environmental factors or when the statistical approach is underpowered resulting in what is referred to as "voodoo correlations" wherein random chance yields a result that is not repeatable. [42, 43] The solution is to expand our horizons to as many different kinds of chemicals as possible, and then to do our statistical homework to make sure we get solid and repeatable results. In the remainder of this article, we focus on using exhaled breath, but the general approaches work just as well for other biological media such as blood and urine.

Biomarkers in exhaled breath condensate (EBC) and aerosols (EBA)

The "Holy Grail" in preventative medicine is diagnosing early disease state before overt symptoms and irreversible effects of the disease occur. Similarly, figuring out which chemicals in the environment might be related to preclinical effects or perturbations in the human metabolism is the ultimate goal for exposure science. Exploiting breath as the biological fluid for analysis (rather than blood or urine) is attractive because sampling is non-invasive, can be repeated frequently, and there is essentially an unlimited supply. Basically, using breath biomarkers means no syringes full of blood, no containers of urine, no disposal of infectious wastes - just collecting a bit of gas and mist that otherwise get lost to the ambient air.

We have found that breath is not just for gas-phase compounds analysis; in addition to dissolved volatile organics, the small water droplets and mists we exhale also contain a myriad of interesting large molecules, proteins, bacteria, and viruses that cannot be measured with traditional gas chromatography.[40] A great deal of the literature revolves around exhaled breath condensate (EBC) as an alternative to breath gas collection, primarily for measuring pH and inflammatory markers as indicators of asthma status for chronic obstructive pulmonary disease and for pulmonary infections.[41-44] EBC is also useful for interpreting the effects of recent exposures to environmental contaminants both for assessing exogenous chemicals and their metabolic responses.[45-47] It is notable that many factors can affect the composition of EBC including the pulmonary microbiome, specific pathologies, diet, epigenesis, and other individualized influences, but these are beyond the scope of this discussion.

Briefly, the exhaled breath is passed through a chilled tube and the pulmonary water vapor and aerosols condense on the surfaces. Typically, a subject breathes through the device for 10 minutes or so and about 1 ml of condensate is subsequently harvested from the tube. There are a variety of devices commercially available; in our laboratory we use the "R-Tube" (Respiratory Research, Austin, TX) modified for dry ice temperature collection.[48] The EBC fluid is a useful medium for a variety of polar volatile organic

compounds, cytokines, and large molecules, but tends to be very dilute as it is more than 99% water.

To collect just the aerosol fraction, we have been implementing a new technique wherein the subject blows through a filter at room temperature; we do not attempt to capture the water vapor or the more volatile fractions of breath. This is an attractive method for a qualitative study because it requires no infrastructure in the field besides disposable filters and sealable plastic bags. Furthermore, the filters collect exhaled pulmonary bacteria that stay alive at room temperature and can be cultured later in the laboratory. In addition to being able to identify strains via standard microbiological techniques, we also collect the headspace of the bacteria for subsequent organics analyses with GC-MS. In the long run, we hope to learn enough about individual bacteria "fingerprints" of gas-phase molecules that we can jump directly into exhaled breath analysis (without off-line bacterial cultures) to discern infection status of human subjects. The whole breath samples could also be collected directly onto adsorbent media (e.g. Tenax© tubes) that can capture both aerosol and gas-phase constituents.

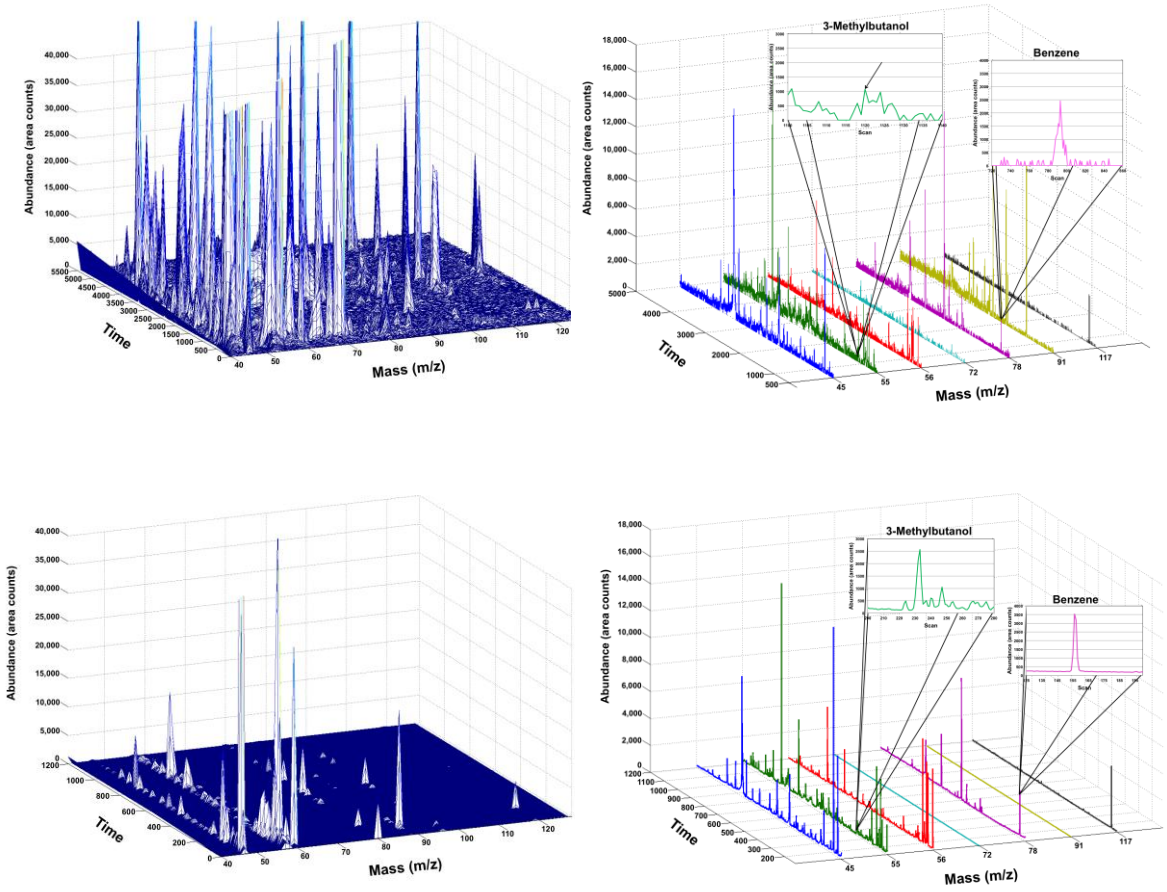
Comparisons of discovery and targeted biomarker data

In the following discussion, we describe the difference between discovery and targeted biomarker research using GC-MS results from exhaled breath samples as an example. These samples are drawn from our archives from methods development projects using anonymous biological specimens. Different sampling and analytical schemes will have similar properties. This is to be considered the simplest of examples using conventional benchtop laboratory instrumentation available in many environmental and clinical laboratories (Ultra TD autosampler coupled with a Unity thermal desorber - Markes International, Ltd, Llantrisandt, UK, Agilent 6890N GC coupled with Agilent 5973I MS - Agilent, Palo Alto, CA).

If we collect whole breath (e.g. onto adsorbent tubes) and run it through a GC-MS under "scan" mode, the instrument essentially looks for everything that can possibly pass through the column and can create an ion in the MS source. Typically, we scan from 33 m/z (mass to charge ratio) to 350 m/z in default 0.1 m/z increments resulting in large data files with reasonable sensitivity at any given m/z. We realize that this resolution does not add analytical advantage over the nominal single unit amu data, but we choose this default setting to draw more understandable graphics. After we run a sample, we can look for benzene, for example, by extracting the 78 m/z ion from all ions and locate that spot in time on the GC run; for ethanol we look for 45 m/z, for 1-propanol we look for 59 m/z; for 2-methylbutanol we look for 55 m/z, for 1-butanol and hexanal we look for 56 m/z, for toluene we look for 91 m/z, for carbon tetrachloride we look for 117, etc.[48] By combining elution time from the column and the specific ions for a molecule, we can dig out peaks and values for most compounds that are present in the breath and can respond on the instrument without previous knowledge. This way, we can "discover" what volatile organic compounds are in a particular breath sample. Certainly we check to see if we are right by also analyzing true standards and matching time and complete mass spectra. We thus create a long list of compounds that potentially come from breath and

begin to understand relative amounts, variability, and relationships depending on subject meta-data (gender, age, ethnicity, socio-economics, location, etc.).

At some point, we no longer want every possible compound but would like to do a more specific and sensitive job for only a few compounds of exceedingly high interest. In this case, we go into SIM (selective ion monitoring) acquisition mode and only acquire the ions for specific molecules. In following the previous scan example, we would "SIM" for only pre-selected ions such as 45, 55, 56, 78, and 117 m/z and thus spend much more time looking for those by sacrificing all of the other ions between 33 and 350 m/z. (We point out that for routine analyses of actual samples, we always include confirmatory ions and test for ion ratios for each target analyte to assure quality of results). SIM is a targeted approach resulting in a much lower background noise level and a higher (or cleaner) response for the chosen molecules. The figure below shows a series of GC-MS plots comparing Scan and SIM GC-MS runs from identical breath samples; we also show the respective extractions for each of the methods and compounds wherein we see that SIM gives better data, but many other compounds are completely gone. We use benzene (78 m/z) and 3-methylbutanol (55 m/z) as examples of ubiquitous exogenous and endogenous compounds, respectively) in breath and further extract the small portion of the chromatogram to highlight the improvement in analytical results going from discovery to targeted analysis.



Left side graphics are a graphic depiction of a "whole breath" gas sample analyzed with GC-MS under full scan, or discovery mode (top), and with selected ion monitoring, or targeted mode (bottom). The right side graphics show only the extracted traces for specific ions, and a close-up for two specific compounds, 3-methylbutanol and benzene. Although the extractions of the discovery mode (top) show a "noisier" result than those of the targeted mode (bottom), the targeted mode is missing many ions that can be seen in discovery mode.

This is only a simple example that contrasts discovery vs. targeted analysis. There are many different analytical systems that can collect either or both types of modes.[49-51] For example, liquid chromatography tandem mass spectrometry (LC-MS-MS) is a powerful tool for making very exact measurements of larger molecules, but we can only take advantage of the instrument's specificity by targeting analytes *a priori* and setting up their specific transitions for the linear quadrupoles.[52] In contrast, liquid chromatography time of flight mass spectrometry (LC-ToF), or LC-Orbitrap instrumentation can be used in discovery mode and the targeted features could be applied in data post-processing.[53-55]

Summary of discovery approaches in breath biomarker research

The use of discovery approaches in all biomarker research (not just in breath) is extremely important in that it allows us to get an overall view of what is going on. Certainly not every compound is relevant for answering a particular question or testing a hypothesis. This requires many samples and analyses from different groups of subjects, and from multiple times from the same subject where we try to look for any changes between or within person that may mean something. From careful observation of such changes, we can build a more specific list of compounds that can help deduce the effects of exposures on health state. There are, however, a number of precautions we need to take during the discovery phase of any study.

The first issue has been discussed above; by virtue of the broad nature of discovery analyses, any individual compound is not optimized for detection and so the data tend to be more qualitative. Also, the molecules with the highest concentrations and occurrence may not be the most important, and so we run the risk of missing something down in the "noise" that could tell a better story.

The second issue revolves around "test - retest" or "multiple hypothesis" statistics. The more compounds we are comparing the more likely we will have false positives (the aforementioned "voodoo correlation"); and the more we restrict the possibility of having excess false positives (e.g. Bonferroni correction), we increase the probability of false negatives. The statistics improve as we increase the number of independent samples and repeat the study with other groups of subjects. Other post-processing methods including co-linearity tests, non-parametric association tests, mixed models, and variable clustering can also reduce the occurrence of false positives but these methods may blur the impact of individual compounds and increase the chance of erroneously discarding an important probative compound.[56-58]

As such, the discovery approach is the first step. Without it, we would only be guessing as to what to look for with our specific targeted approaches. However, we cannot become complacent and assume that discovery is all-inclusive. We suggest that multiple discovery experiments with different biological media and analytical platforms are the ultimate solution for finding as many probative biochemicals as possible. Only then can we streamline the targeted approaches to achieve high throughput sample processing. We conclude that exposure science is indeed evolving along a path blazed by the genomics community by embracing the broad discovery aspects of biomarker chemicals and defining the complex human exposome in a top down fashion, yet still recognizing the value of the targeted approaches that lead to exposure mitigation and regulation for protecting public health.

ACKNOWLEDGEMENTS

We are indebted to Terence Risby from Johns Hopkins University, Stephen Rappaport from University of California, Berkeley, and Stephen Edwards, Jon Sobus, Myriam Medina-Vera, Michael Madden, Andrew Ghio, and Linda Sheldon from U.S. EPA for invaluable insights and discussions over the past few years. We also thank our many colleagues in the exposure science and breath diagnostics/analysis communities who provide inspiration for continuing to develop the science for protecting human health. We especially thank C. Gaul and K. Tarpley from SRA International CreativeTeam for producing some of the artwork. This work was reviewed by the U.S. EPA and approved for publication. The views expressed in this article are those of the author[s] and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. The authors declare they have no competing financial interests.

AUTHOR INFORMATION

Joachim D. Pleil is principal investigator for systems biology approaches for exposure assessment for the U.S. EPA's Office of Research and Development in Research Triangle Park, NC and serves as adjunct Professor of Environmental Sciences and Engineering at the University of North Carolina School of Public Health in Chapel Hill, NC. He is a founding member of the International Association of Breath Research (IABR); he serves on the science advisory board of IABR and on the editorial boards of the Institute of Physics (IOP) Journal of Breath Research (JBR) and the InformaHealthCare journal Biomarkers. Matthew A. Stiegel is an ORISE Research Fellow at U.S. EPA and a Ph.D. candidate in the Department of Environmental Sciences and Engineering at the University of North Carolina School of Public Health in Chapel Hill, NC.

REFERENCES

(1) Pleil, J. D.; Blount, B. C.; Waidyanatha, S.; Harper, M. *J. Expo. Sci. Environ. Epidemiol.* **2012**, *22*, 317-9.
(2) Barr, D. B. *J. Expo. Sci. Environ. Epidemiol.* **2006**, *16*.
(3) Lioy, P. J. *Environ. Health. Perspect.* **2010**, *118*, 1081-90.
(4) Cohen-Hubal, E. A.; Barr, D. B.; Koch, H. M.; Bahadori, T.; J. *J. Expo. Sci. Environ. Epidemiol.* **2011**, *21*, 121-122.
(5) Wild, C. P. *Cancer. Epidemiol. Biomarkers. Prev.* **2005**, *14*, 1847-50.
(6) Pleil, J. D.; Stiegel, M. A.; Madden, M. C.; Sobus, J. R. *Chemosphere* **2011**, *84*, 716-23.
(7) Helix Project: Building the Early-Life Exposome. http://www.projecthelix.eu/ (accessed(9/10/2013).
(8) National Research Council of The National Academies. The National Academies Press, Washington, D.C., **2013**.
(9) Berkeley Center for Exposure Biology: University of California, Berkeley, Biological Response Indicators of Environment Stress Centers. http://www.niehs.nih.gov/research/

supported/dert/cris/programs/exposure/biological_response/berkeley/index.cfm (accessed 9/10/2013).
(10) Health and Exposome Research Center: Understanding Lifetime Exposures. http://news.emory.edu/stories/2013/05/hercules_esposome_grant_rollins/(accessed 9/10/2013).
(11) Rappaport, S. M.; Smith, M. T. *Science* **2010**, *330*, 460-1.
(12) Pleil, J. D.; Sheldon, L. S. *Biomarkers.* **2011**, *16*, 99-105.
(13) Pleil, J. D. *J. Toxicol. Environ. Health. B. Crit. Rev.* **2012**, *15*, 264-80.
(14) Paoloni-Giacobino, A. *Swiss Med. Wkly.* **2011**, *141*, w13321.
(15) Klepeis, N. E.; Nelson, W. C.; Ott, W. R.; Robinson, J. P.; Tsang, A. M.; Switzer, P.; Behar, J. V.; Hern, S. C.; Engelmann, W. H. *J. Expo. Anal. Environ. Epidemiol.* **2001**, *11*, 231-52.
(16) Zhang, L.; McHale, C. M.; Rothman, N.; Li, G.; Ji, Z.; Vermeulen, R.; Hubbard, A. E.; Ren, X.; Shen, M.; Rappaport, S. M.; North, M.; Skibola, C. F.; Yin, S.; Vulpe, C.; Chanock, S. J.; Smith, M. T.; Lan, Q. *Chem. Biol. Interact.* **2010**, *184*, 86-93.
(17) Rappaport, S. M. *J. Expo. Sci. Environ. Epidemiol.* **2011**, *21*, 5-9.
(18) Liu, L.; Li, Y.; Tollefsbol, T. O. *Curr. Issues. Mol. Biol.* **2008**, *10*, 25-36.
(19) The White House, Office of the Press Secretary, PRESIDENT CLINTON ANNOUNCES THE COMPLETION OF THE FIRST SURVEY OF THE ENTIRE HUMAN GENOME Hails Public and Private Efforts Leading to This Historic Achievement; http://web.ornl.gov/sci/techresources/Human_Genome/project/clinton1.shtml (accessed 9/10/2013).
(20) Wade, N. In *New York Times*; New York Times: New York, NY, **2010**.
(21) Hall, S. S. *Sci. Am.* **2010**, *303*, 60-7.
(22) Kaiser, J. *Science (New York, N.Y.)* **2012**, *338*, 1016-7.
(23) Egeghy, P. P.; Hauf-Cabalo, L.; Gibson, R.; Rappaport, S. M. *Occup. Environ. Med.* **2003**, *60*, 969-76.
(24) Rigas, M. L.; Okino, M. S.; Quackenboss, J. J. *Toxico. Sci.* **2001**, *61*, 374-81.
(25) Toxic Substances Control Act. *U.S. Code*, Chapter 53, Title 15, 2601-2629, **1976**.
(26) U.S. Environmental Protection Agency. *Framework for an EPA Chemical Safety for Sustainability Research Program*: Washington, D. C., **2011**.
(27) Judson, R.; Richard, A.; Dix, D. J.; Houck, K.; Martin, M.; Kavlock, R.; Dellarco, V.; Henry, T.; Holderman, T.; Sayre, P.; Tan, S.; Carpenter, T.; Smith, E. *Environ. Health. Perspect.* **2009**, *117*, 685-95.
(28) Sheldon, L. S.; Cohen Hubal, E. A. *Environ. Health. Perspect.* **2009**, *117*, 119-1194.
(29) National Research Council of The National Academies, Toxicity Testing in the 21st Century: A Vision and a Strategy; **2007**.
(30) Dix, D. J.; Houck, K. A.; Martin, M. T.; Richard, A. M.; Setzer, R. W.; Kavlock, R. *Toxico. Sci.* **2007**, *95*, 5-12.
(31) Knudsen, T.; Martin, M.; Chandler, K.; Kleinstreuer, N.; Judson, R.; Sipes, N. *Methods. Mol. Biol.* **2013**, *947*, 343-74.
(32) Bouhifd, M.; Hartung, T.; Hogberg, H. T.; Kleensang, A.; Zhao, L. *J. Appl. Toxicol.* **2013**. doi:10.1002/jat.2874.
(33) Pleil, J. D.; Williams, M. A.; Sobus, J. R. *Toxicol. Lett.* **2012**, *215*, 201-7.

(34) Ankley, G. T.; Bennett, R. S.; Erickson, R. J.; Hoff, D. J.; Hornung, M. W.; Johnson, R. D.; Mount, D. R.; Nichols, J. W.; Russom, C. L.; Schmieder, P. K.; Serrrano, J. A.; Tietge, J. E.; Villeneuve, D. L. *Environ. Toxicol. Chem.* **2010**, *29*, 730-41.
(35) Liu, H.; Bebu, I.; Li, X. *Front. Biosci.* **2010**, *2*, 325-38.
(36) Patel, C. J.; Bhattacharya, J.; Butte, A. J. *PloS. One.* **2010**, *5*, e10746.
(37) Rappaport, S. M. *Biomarkers.***2012**, *17*, 483-9.
(38) Amann, A.; Corradi, M.; Mazzone, P.; Mutti, A. *Expert. Rev. Mol. Diag.* **2011**, *11*, 202-217.
(39) Mazzone, P. J. *J. Breath. Res.* **2012**, *6*, 027106.
(40) Slatore, C. G.; Gould, M. K.; Au, D. H.; Deffebach, M. E.; White, E. *BMC. Cancer.* **2011**, *11*, 7.
(41) Bunn, P. A. *Arch. Pathol. Lab. Med.* **2012**, *136*, 1478-81.
(42) Pleil, J. D.; Stiegel, M. A.; Risby, T. H. *J. Breath. Res.* **2013**, *7*, 017107.
(43) Miekisch, W.; Herbig, J.; Schubert, J.K. *J. Breath. Res.* **2012**, *6*, 036007.
(44) Rosias, P. *J. Breath. Res.* **2012**, *6*, 027102.
(45) Loukides, S.; Kontogianni, K.; Hillas, G.; Horvath, I. *Curr. Med. Chem.* **2011**, *18*, 1432-43.
(46) Kazani, S.; Israel, E. *J. Breath. Res.* **2010**, *4*, 047001.
(47) Lee, W.; Thomas, P. S. *Clin. Transl. Sci.* **2009**, *2*, 150-5.
(48) Chambers, S. T.; Scott-Thomas, A.; Epton, M. *Curr. Opin. Pulm. Med.* **2012**, *18*, 228-32.
(49) Zhang, J.; Zhu, T.; Kipen, H.; Wang, G.; Huang, W.; Rich, D.; Zhu, P.; Wang, Y.; Lu, S.-E.; Ohman-Strickland, P.; Diehl, S.; Hu, M.; Tong, J.; Gong, J.; Thomas, D.; Committee, H. E. I. H. R. *Res. Resp. Health. Eff. Inst.* **2013**, 5-174.
(50) Hubbard, H. F.; Sobus, J. R.; Pleil, J. D.; Madden, M. C.; Tabucchi, S. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **2009**, *877*, 3652-8.
(51) Sawyer, K.; Samet, J. D.; Ghio, A. J.; Pleil, J. D.; Madden, M. C. *J. Breath. Res.* **2009**, *2*, 037019.
(52) Pleil, J. D.; Hubbard, H. F.; Sobus, J. R.; Sawyer, K.; Madden, M. C. *J. Breath. Res.* **2008**, *2*, 026001.
(53) Koester, C. J.; Moulik, A. *Anal. Chem.* **2005**, *77*, 3737-54.
(54) Boja, E. S.; Rodriguez, H. *Proteonomics* **2012**, *12*, 1093-110.
(55) Dasilva, N.; Díez, P.; Matarraz, S.; González-González, M.; Paradinas, S.; Orfao, A.; Fuentes, M. *Sensors (Basel)* **2012**, *12*, 2284-308.
(56) Grebe, S. K.; Singh, R. J. *Clin. Biochem. Rev.* **2011**, *32*, 5-31.
(57) Lacorte, S.; Fernandez-Alba, A. R. *Mass. Spectrum. Rev.* **2006**, *25*, 866-80.
(58) Ferrer, I.; Thurman, E. M.; Fernández-Alba, A. R. *Anal. Chem.* **2005**, *77*, 2818-25.
(59) Zubarev, R. A.; Makarov, A. *Anal. Chem.* **2013**, *85*, 5288-96.
(60) Sobus, J. R.; Pleil, J. D.; McClean, M. D.; Herrick, R. F.; Rappaport, S. M. *Toxico. Lett.* **2010**, *199*, 247-53.
(61) Pleil, J. D.; Stiegel, M. A.; Sobus, J. R. *J. Breath. Res.* **2011**, *5*, 046005.
(62) Pleil, J. D.; Sobus, J. R.; Sheppard, P. R.; Ridenour, G.; Witten, M. L. *Chem. Biol. Interact.* **2012**, *196*, 68-78.