Toxicology Reports 3 (2016) 29–45

Contents lists available at ScienceDirect

Toxicology Reports

journal homepage: www.elsevier.com/locate/toxrep

Sequencing the exposome: A call to action☆

Dean P. Jones*

Department of Medicine, Emory University, Whitehead Biomedical Research Building, Suite 205P, 615 Michael Street, NE Atlanta, GA 30322, USA

ARTICLE INFO

Article history:
Received 25 November 2015
Accepted 25 November 2015
Available online 28 November 2015

Keywords:
Mass spectrometry
Biomonitoring
Analytical chemistry
Metabolomics
Environmental surveillance

ABSTRACT

The exposome is a complement to the genome that includes non-genetic causes of disease. Multiple definitions are available, with salient points being global inclusion of exposures and behaviors, and cumulative integration of associated biologic responses. As such, the concept is both refreshingly simple and dauntingly complex. This article reviews high-resolution metabolomics (HRM) as an affordable approach to routinely analyze samples for a broad spectrum of environmental chemicals and biologic responses. HRM has been successfully used in multiple exposome research paradigms and is suitable to implement in a prototype universal exposure surveillance system. Development of such a structure for systematic monitoring of environmental exposures is an important step toward sequencing the exposome because it builds upon successes of exposure science, naturally connects external exposure to body burden and partitions the exposome into workable components. Practical results would be repositories of quantitative data on chemicals according to geography and biology. This would support new opportunities for environmental health analysis and predictive modeling. Complementary approaches to hasten development of exposome theory and associated biologic response networks could include experimental studies with model systems, analysis of archival samples from longitudinal studies with outcome data and study of relatively short-lived animals, such as household pets (dogs and cats) and non-human primates (common marmoset). International investment and cooperation to sequence the human exposome will advance scientific knowledge and also provide an important foundation to control adverse environmental exposures to sustain healthy living spaces and improve prediction and management of disease.

1. Introduction and definitions

The exposome is the cumulative measure of environmental influences and biological responses throughout the lifespan [28]. The definition is inclusive, consistent with Christopher Wild's concept [44] to broadly complement the genome with non-genetic factors impacting human health. Measurements of all environmental, dietary, microbiome, behavioral, therapeutic and endogenous processes present a daunting challenge for systematic study, especially when considered cumulatively throughout life.

Less than a half-century ago, sequencing the human genome was imaginable by few scientists. Yet visionary leaders championed this goal, societies and industries invested in new technologies, and success was attained. Despite monumental barriers, earlier generations succeeded in quests to control communicable diseases, eradicate smallpox, prevent polio, cure childhood leukemia, etc. Sequencing the human exposome is a formidable challenge for contemporary science and technology, but with vision and commitment, this is attainable.

Wild [44] provided a thoughtful outline to address the challenge, and others have emphasized the need for critical thought and design to pursue this goal [34,1,45]. In the present article, I provide a brief review on environmental applications of high-resolution metabolomics (HRM; see Table 1 for acronym definitions) and comment on the potential to use this as a central framework to initiate sequencing the human exposome. I use the term "sequencing" to emphasize the time domain of the exposome; the cumulative measure of exposures for an individual cannot be sequenced in the same way as the genome can be sequenced. But there are many types of exposure memory systems [17], and there is little doubt that improved understanding of epigenetics and other exposure memory systems will allow certain aspects of an individual's exposome to be measured retrospectively. So the present review of HRM is not intended as a final chapter on how to sequence the exposome, but rather as a place to begin.

I first describe the use of high-resolution mass spectrometry for clinical metabolic profiling and discuss advantages and

☆ This article was presented in a session entitled "Genome, Epigenome and Exposome" at the 45th Annual Symposium of the Society of Toxicology of Canada, Ottawa, December 4, 2013.

* Corresponding author.
E-mail address: dpjones@emory.edu

http://dx.doi.org/10.1016/j.toxrep.2015.11.009

limitations. This is followed by a discussion of computational workflows, which support both targeted and discovery analyses using commonly accessible biological samples and advanced informatics methods. I then consider use of this analytical platform as a universal surveillance tool to monitor environmental exposures and evaluate biological effects. Finally, I discuss HRM as a possible central element of a "Human Exposome Project" to sequence the cumulative environmental influences and biological responses throughout lifespan. This is presented as a call to the international environmental health research community to champion this effort and work together in this common goal. At Emory University, in collaboration with Georgia Institute of Technology and with support of the National Institute of Environmental Health Sciences, we are building toward this goal through the HERCULES Exposome Research Center directed by Gary W. Miller, Ph.D. (http://humanexposomeproject.com).

2. Rationale for development of high-resolution metabolomics

Analytical traditions, as well as regulatory and policy needs of government, have limited flexibility to address important challenges in health and environmental research. During the period from 1958, when automated amino acid analysis was introduced [29], and 2007, when we began experimental development of advanced blood chemistry analyses with high-resolution mass spectrometry [16], analytical chemistry provided merely a ten-fold improvement in the number of chemicals that could be measured in a routine analysis of plasma or serum, i.e., from about 30 to 300. During this time period, DNA sequencing progressed from a complete inability to sequence the human genome to ability to accomplish the task within a few days.

While considerable financial investment through the Human Genome Project contributed importantly to this success, practical differences from analytical chemistry also existed in the scope of the human genome initiative and in the tolerance of the project for errors. Genome sequencing was intended to be more or less complete and was pursued even for genes without known function. Additionally, whole genome sequencing was developed with an expectation for errors in sequencing and assembly. Instead of abandonment because methods were inadequate or errors were common, progress was made by embracing new methods and developing new approaches to address errors. Although imperfect in many ways, this resulted in overall success.

In contrast, progress in analytical coverage of small molecules in biologic systems appears to have languished more due to analytical traditions than to limitations in technology. With a focus on matching proteomic capabilities to those of genomics, tremendous advances in technology were achieved in mass spectrometry (MS). A byproduct of the successes of proteomics was the transformation of MS capabilities for detection and measurement of small molecules.

2.1. Mass spectrometry for chemical profiling

Mass spectrometry (MS) has a special role in analytical chemistry because the mass of a chemical is an absolute property. Thus, if the measured mass does not match that of the purported chemical, then the chemical identification is incorrect. MS involves measurement of chemicals or derived fragments of chemicals as ions (m/z, mass-to-charge ratio) in the gas phase [13]. The ions can be formed by interaction of a neutral chemical with H^+, Na^+ or other cation, by loss of H^+ as occurs with ionization of carboxylic acids, or by dissociation of a chemical into product ions. With introduction of electrospray ionization (ESI) [47], routine measurement of a very broad range of small molecules in biological materials without extensive fragmentation became practical.

Mass spectrometers have three components, an ion source, which generates ions; a mass analyzer, which separates ions according to m/z; and a mass detector, which measures ion intensity, i.e., the amount of ions with the respective m/z (Fig 1A). Many physical and chemical principles are used for MS instruments, and a recent overview is available [25]. Mass spectrometers differ in sensitivity to detect low abundance ions, resolve ions with very similar m/z, and provide accurate estimates of m/z. For instruments with poorer mass resolution and mass accuracy, chromatographic separation prior to MS is necessary to allow measurement of chemicals with the same unit mass (Fig 1B). Both gas chromatography (GC) and liquid chromatography (LC) are often coupled to MS to improve measurement of chemicals according to characteristic retention time and m/z (Fig 1A and B).

Quadrupole mass analyzers, often designated simply by "Q", selectively stabilize or destabilize the paths of ions passing through oscillating electrical fields between 4 parallel rods. Common

Table 1
Abbreviations and acronym definitions [For additional MS details, see [30]].

Term	Definition	Term	Definition
AMU	Atomic mass unit	m/z	Mass to charge ratio for an ion measured by mass spectrometry
apLCMS	Adaptive processing algorithms for extraction of MS data	MB	Maneb, a fungicide
Asp	Aspartate	MS	Mass spectrometry
Cd	Cadmium	MS/MS	First product ion spectrum from ion dissociation MS; same as MS^2
CV	Coefficient of variation	MS^1	Spectrum of a precursor ion
ESI	Electrospray ionization	MS^2	First product ion spectrum from ion dissociation MS; same as MS/MS
FDR	False discovery rate	MS^n	Product ion spectra from a sequence of ion dissociations
FT	Fourier transform	MWAS	Metabolome-wide association study
FT-ICR	Type of mass analyzer	PBDE	Polybrominated diphenyl ether,
GC	Gas chromatography	PD	Parkinson disease
GC–MS	Combined gas chromatography and mass spectrometry	Phe	Phenylalanine
GWAS	Genome-wide association study	PLS-DA	Partial least squares-discriminant analysis
HIV	Human immunodeficiency virus	PQ	Paraquat, an herbicide
HMDB	Human metabolome database	Q	Quadrupole, a type of mass analyzer
HPLC	High performance liquid chromatography	Q-TOF	Tandem mass spectrometer combining quadrupole and TOF
HRM	High-resolution metabolomics	SOP	Standard operating procedures
ICR	Ion cyclotron resonance, a type of mass analyzer	SRM1950	National institute of standards pooled reference human plasma
Ile	Isoleucine	TMWAS	Transcriptome metabolome-wide association study
KEGG	Kyoto encyclopedia of genes and genomes	TOF	Time-Of-Flight, a type of mass analyzer
LC	Liquid chromatography	Tyr	Tyrosine
LC-MS	Combined liquid chromatography and mass spectrometry	UHRAM	Ultra-high resolution accurate mass, used in reference to FT MS
Leu	Leucine	XCMS	Algorithms for MS data extraction
LIMMA	Linear models for microarray, software for differential expression	xMSanalyzer	Algorithms to improve data extraction by apLCMS or XCMS