Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemico-Biological Interactions 153–154 (2005) 123–127
Chemico-Biological Interactions
www.elsevier.com/locate/chembioint

Use of 'Omic' technologies to study humans exposed to benzene

Martyn T. Smith [a,*], Roel Vermeulen [b], Guilan Li [c], Luoping Zhang [a], Qing Lan [b], Alan E. Hubbard [a], Matthew S. Forrest [a], Cliona McHale [a], Xin Zhao [a], Laura Gunn [a], Min Shen [b], Stephen M. Rappaport [d], Songnian Yin [c], Stephen Chanock [b], Nathaniel Rothman [b]

[a] *School of Public Health, 140 Warren Hall, University of California, Berkeley, CA 94720-7360, USA*
[b] *Division of Cancer Epidemiology and Genetics, NCI, NIH, DHHS, Bethesda, MD 20892, USA*
[c] *Chinese Center for Disease Control and Prevention, Beijing, China*
[d] *School of Public Health, University of North Carolina, Chapel Hill, NC 27599, USA*

Available online 19 April 2005

Abstract

'Omic' technologies include genomics, transcriptomics (gene expression profiling), proteomics and metabolomics. We are utilizing these new technologies in an effort to develop novel biomarkers of exposure, susceptibility and response to benzene. Advances in genomics allow one to study hundreds to thousands of single nucleotide polymorphisms simultaneously on small quantities of DNA using array-based technologies. We are currently utilizing these technologies to examine genetic variation in pathways relating to biotransformation, DNA repair, folate metabolism and immune response with the goal of finding biomarkers of susceptibility to benzene hematotoxicity. Transcriptomics is used to measure the full complement of activated genes, mRNAs or transcripts in a particular tissue at a particular time typically using microarray technology. We have applied microarrays to the study of global gene expression in the peripheral blood cells of benzene-exposed workers. More than 100 genes were identified as being potentially differentially expressed, with genes related to apoptosis and immune function being the most significantly affected. Initial studies employing proteomics have also shown that several proteins are altered in the serum of exposed compared to control subjects and these proteins are potential biomarkers of benzene exposure. Omic technologies therefore have significant potential in generating novel biomarkers of exposure, susceptibility and response to benzene.

Keywords: Benzene; Biomarkers; Genomics; Proteomics; Microarray

1. Introduction

'Omics' is a general term used to describe several rapidly growing fields of scientific endeavor, the best known member of which is genomics [1]. Genomics is

* Corresponding author. Tel.: +1 510 642 8770; fax: +1 510 642 0427.
E-mail address: martynts@berkeley.edu (M.T. Smith).

0009-2797/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cbi.2005.03.017

the study of a genome, the complete genetic complement of an individual or species, rather than the study of single genes. The suffix-omics generally refers to the study of a complete set of biological molecules. Just as genomics is the study of an organisms genome, proteomics is the study of an organisms entire complement of proteins and metabolomics is the study of the complete set of low-molecular weight metabolites present in a cell or organism at any one time [2,3]. These 'omic' technologies allow for a large number of endpoints to be simultaneously measured on biological samples from human and animal subjects [4,5]. Their application in toxicology and molecular epidemiology holds great promise, but it should be realized that these new technologies are discovery tools rather than traditional assays and, as such, data generated by them should be treated with caution. We are currently utilizing these new 'omic' technologies, in an effort to discover novel biomarkers of exposure, susceptibility and response to benzene.

2. Using genomics to study susceptibility to benzene toxicity

Advances in genomics allow for a large percentage of human genetic variation to be studied in subjects participating in epidemiological studies. For example, it is now possible to study thousands of single nucleotide polymorphisms simultaneously on small quantities of DNA isolated from blood, buccal cells or other bodily fluids using array-based technologies from Illumina, Affymetrix and ParAllele [6–9]. This allows for functional variant and haplotype information to be obtained on hundreds of genes simultaneously. To ameliorate statistical issues arising from multiple comparisons, current approaches focus mainly on hypothesis-driven studies of genes in particular pathways rather than whole genome scans. For example, we are examining genetic variation in pathways relating to DNA repair, folate metabolism and immune response to determine their role in susceptibility to benzene hematotoxicity. Previous studies of variants in genes related to the metabolic activation of benzene have shown that several of these confer susceptibility to hematotoxicity [10–13], indicating that this should also be studied more thoroughly with genomic technology.

3. Gene expression profiling of benzene exposure with transcriptomics

Transcriptomics is used to measure the full complement of activated genes, mRNAs or transcripts in a particular tissue at a particular time typically using microarray technology [1,14]. Microarrays use immobilised cDNA or oligonucleotide probes to simultaneously monitor the expression of thousands of genes and obtain a view of global gene expression (i.e. a view of all mRNA transcripts expressed by a cell = the transcriptome) [15]. These have recently been used to investigate variation in gene expression in the peripheral blood leukocytes of normal individuals [16] and are becoming increasingly used in toxicology and molecular epidemiology [5,17,18]. We hypothesized that microarrays could be used to identify changes in gene expression caused by human exposure to benzene and provide information on mechanisms of benzene toxicity. We have therefore applied microarrays to the study of global gene expression in the peripheral blood cells of benzene-exposed workers [19].

A potential problem with employing microarrays in molecular epidemiology studies is the fact that mRNA is unstable [20]. We overcome this problem by performing the first step of RNA isolation in the field and stabilizing the RNA for later analysis. We analyzed RNA stabilized in this fashion from benzene-exposed workers and matched controls using the most comprehensive and standardized human array available at the time, the U133 Affymetrix GeneChip [21]. U133 chips together contain almost 45,000 probe sets, representing >39,000 unique transcripts derived from ~33,000 well substantiated human genes allowing one to obtain a global view of gene expression. We identified more than one hundred genes as being potentially differentially expressed, with genes related to apoptosis and immune function being the most significantly affected by benzene exposure [19]. Thus, microarray analysis revealed that global gene expression was altered in peripheral blood cells following benzene exposure and that alterations in the expression of certain specific genes may be useful biomarkers of early effect. Our studies in humans compliment other toxicogenomic studies in experimental animals and in vitro systems published previously [17,22], and discussed elsewhere in this volume.

4. Application of proteomics in benzene biomarker research

In light of the fact that the human genome consists of approximately 20–25,000 protein-coding genes, a fraction of what was originally expected, it has become clear that mammalian systems are more complex than genes alone [23]. Alternate splicing, as well as over 200 post-translational modifications affect a proteins structure, function, stability and interactions with other molecules [24]. A number of different proteins are therefore likely to be expressed by a single gene. Proteomics studies all of an organisms proteins or proteome. This can contain thousands of proteins, which, even within a given organism or cell, can vary depending on cell or tissue type, disease state and other factors. To decipher a proteome, proteomics begins with the systematic separation and identification of all proteins within a cell, tissue, or other biological sample [2]. Although proteomics has traditionally focused on quantitative analysis of protein expression, more recently, it has expanded to include structural analysis and identification of proteins. Application of proteomic technologies to benzene research could provide new biomarkers that reflect functional changes in protein expression induced by exposure and causal changes related to toxicity and cancer induction.

Proteomic research has recently been adapted to high throughput, highly sensitive technologies, however, much is based on methodology that has been used for decades. Two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS) have been and are still used today for the analysis and identification of differentially expressed proteins. However, new technology, particularly protein arrays have increased throughput, sensitivity and may overcome many of the limitations of traditional proteomic methods [25].

Joo et al. [26] recently analyzed the proteins in plasma of workers exposed to low levels of benzene by 2-DE. Fifty workers from a printing company and 38 matched unexposed healthy subjects were enrolled in the study. The mean values of trans, trans-muconic acid in workers exposed to benzene and in unexposed subjects were 1.011 +/− 0.249 and 0.026 +/− 0.028 mg/g creatinine, respectively. Protein profiles were significantly different ($P < 0.05$) in the two groups, as identified by matrix-assisted laser desorption ionization/time of flight (MALDI–TOF) mass spectrometry and confirmed by Western blot. T cell receptor beta chain (TCR beta), FK506-binding protein (FKBP51) and matrix metalloproteinase-13 (MMP13) were found to be up-regulated in the benzene-exposed workers. The authors concluded that TCR beta in plasma could be used for the early detection of exposure to benzene. This requires confirmation in other studies of benzene-exposed workers.

Despite its utility, 2-DE is not sensitive enough for low abundance proteins, which may be masked by the presence of high abundance proteins, or for the detection of low-molecular weight or membrane proteins, highly acidic or alkaline proteins. These limitations can be overcome by emerging technologies, particularly surface enhanced laser desorption/ionisation (SELDI), which combines chip technology and MS [25]. SELDI based protein chips utilize affinity capture to bind proteins with unique biochemical properties which allow the rapid capture, purification and characterization of protein directly on the chip. The SELDI process does not require sample purification or labeling. A crude sample may be applied directly to the chip, followed by washes to remove unbound protein, etc [27,28]. The laser is then applied to the bound sample directly on the chip, which then sends ions through the time of flight tube and is thus called SELDI–TOF–MS. A mass spectrum is generated, and may be compared with available databases.

In recent studies by our group, serum samples from benzene-exposed workers and matched controls were subjected to anion exchange fractionation and bound to Ciphergen ProteinChip® Arrays. Protein expression patterns were detected by SELDI–TOF–MS as described above. We found that several proteins were consistently altered in exposed compared to control subjects (unpublished data). Identification of these proteins is currently underway.

5. Conclusion

Omic technologies have significant potential in generating novel biomarkers of exposure, susceptibility and response to benzene. Several genetic polymorphisms have already been linked with susceptibility to benzene toxicity and genomic approaches are being used to explore this issue further. Gene expression profiling or transcriptomics may identify novel patterns of

gene expression associated with benzene exposure and provide insights into the mechanism of benzene toxicity. Initial studies employing proteomics in our lab and others have also shown that several proteins are altered in the serum of exposed compared to control subjects and these proteins are potential biomarkers of benzene exposure and response.

Acknowledgements

Supported in part by NIH grants RO1ES06721, P30ES01896 and P42ES04705 to MTS.

References

[1] M.D. Waters, J.M. Fostel, Toxicogenomics and systems toxicology: aims and prospects, Nat. Rev. Genet. 5 (2004) 936–948.
[2] C.L. de Hoog, M. Mann, Proteomics, Annu. Rev. Genom. Hum. Genet. 5 (2004) 267–293.
[3] P.D. Whitfield, A.J. German, P.J. Noble, Metabolomics: an emerging post-genomic tool for nutrition, Br. J. Nutr. 92 (2004) 549–555.
[4] S. Hanash, Integrated global profiling of cancer, Nat. Rev. Cancer 4 (2004) 638–644.
[5] M.D. Waters, J.K. Selkirk, K. Olden, The impact of new technologies on human population studies, Mutat. Res. 544 (2003) 349–360.
[6] S.J. Sawcer, M. Maranian, S. Singlehurst, T. Yeo, A. Compston, M.J. Daly, P.L. De Jager, S. Gabriel, D.A. Hafler, A.J. Ivinson, E.S. Lander, J.D. Rioux, E. Walsh, S.G. Gregory, S. Schmidt, M.A. Pericak-Vance, L. Barcellos, S.L. Hauser, J.R. Oksenberg, S.J. Kenealy, J.L. Haines, Enhancing linkage analysis of complex disorders: an evaluation of high-density genotyping, Hum. Mol. Genet. 13 (2004) 1943–1949.
[7] H. Fakhrai-Rad, J. Zheng, T.D. Willis, K. Wong, K. Suyenaga, M. Moorhead, J. Eberle, Y.R. Thorstenson, T. Jones, R.W. Davis, E. Namsaraev, M. Faham, SNP discovery in pooled samples with mismatch repair detection, Genome Res. 14 (2004) 1404–1412.
[8] D.L. Barker, M.S. Hansen, A.F. Faruqi, D. Giannola, O.R. Irsula, R.S. Lasken, M. Latterich, V. Makarov, A. Oliphant, J.H. Pinter, R. Shen, I. Sleptsova, W. Ziehler, E. Lai, Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel, Genome Res. 14 (2004) 901–907.
[9] A. Oliphant, D.L. Barker, J.R. Stuelpnagel, M.S. Chee, BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping, Biotechniq. Suppl. 56–58 (2002) 60–61.
[10] Q. Lan, L. Zhang, G. Li, R. Vermeulen, R.S. Weinberg, M. Dosemeci, S.M. Rappaport, M. Shen, B.P. Alter, Y. Wu, W. Kopp, S. Waidyanatha, C. Rabkin, W. Guo, S. Chanock, R.B. Hayes, M. Linet, S. Kim, S. Yin, N. Rothman, M.T. Smith, Hematotoxicity in workers exposed to low levels of benzene, Science 306 (2004) 1774–1776.
[11] N. Rothman, M.T. Smith, R.B. Hayes, R.D. Traver, B. Hoener, S. Campleman, G.L. Li, M. Dosemeci, M. Linet, L. Zhang, L. Xi, S. Wacholder, W. Lu, K.B. Meyer, N. Titenko-Holland, J.T. Stewart, S. Yin, D. Ross, Benzene poisoning, a risk factor for hematological malignancy, is associated with the NQO1 609C → T mutation and rapid fractional excretion of chlorzoxazone, Cancer Res. 57 (1997) 2839–2842.
[12] J.N. Xu, C.L. Wu, Y. Chen, Q.K. Wang, G.L. Li, Z. Su, Effect of the polymorphism of myeloperoxidase gene on the risk of benzene poisoning, Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi 21 (2003) 86–89.
[13] J. Wan, J. Shi, L. Hui, D. Wu, X. Jin, N. Zhao, W. Huang, Z. Xia, G. Hu, Association of genetic polymorphisms in CYP2E1, MPO, NQO1, GSTM1 and GSTT1 genes with benzene poisoning, Environ. Health Perspect. 110 (2002) 1213–1218.
[14] S.V. Chittur, DNA microarrays: tools for the 21st century, Comb. Chem. High Throughput Screen 7 (2004) 531–537.
[15] L.M. Staudt, Molecular diagnosis of the hematologic cancers, N. Engl. J. Med. 348 (2003) 1777–1785.
[16] A.R. Whitney, M. Diehn, S.J. Popper, A.A. Alizadeh, J.C. Boldrick, D.A. Relman, P.O. Brown, Individuality and variation in gene expression patterns in human blood, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 1896–1901.
[17] B. Faiola, E.S. Fuller, V.A. Wong, L. Recio, Gene expression profile in bone marrow and hematopoietic stem cells in mice exposed to inhaled benzene, Mutat. Res. 549 (2004) 195–212.
[18] M.M. Wu, H.Y. Chiou, I.C. Ho, C.J. Chen, T.C. Lee, Gene expression of inflammatory molecules in circulating lymphocytes from arsenic-exposed human subjects, Environ. Health Perspect. 111 (2003) 1429–1438.
[19] M.S. Forrest, Q. Lan, A.E. Hubbard, L. Zhang, R. Vermeulen, X. Zhao, Y.-Y. Wu, M. Shen, S. Yin, S. Chanock, N. Rothman, M.T. Smith, Discovery of novel biomarkers by microarray analysis of peripheral blood mononuclear cell expression in benzene-exposed workers, Environ. Health Perspect., in press [online 14 March 2005].
[20] D.C. Thach, B. Lin, E. Walter, R. Kruzelock, R.K. Rowley, C. Tibbetts, D.A. Stenger, Assessment of two methods for handling blood in collection tubes with RNA stabilizing agent for surveillance of gene expression profiles with high density microarrays, J. Immunol. Meth. 283 (2003) 269–279.
[21] C.A. Iacobuzio-Donahue, R. Ashfaq, A. Maitra, N.V. Adsay, G.L. Shen-Ong, K. Berg, M.A. Hollingsworth, J.L. Cameron, C.J. Yeo, S.E. Kern, M. Goggins, R.H. Hruban, Highly expressed genes in pancreatic ductal adenocarcinomas: a comprehensive characterization and comparison of the transcription profiles obtained from three major technologies, Cancer Res. 63 (2003) 8614–8622.
[22] B.I. Yoon, G.X. Li, K. Kitada, Y. Kawasaki, K. Igarashi, Y. Kodama, T. Inoue, K. Kobayashi, J. Kanno, D.Y. Kim, Y. Hirabayashi, Mechanisms of benzene-induced hematotoxicity and leukemogenicity: cDNA microarray analyses using mouse bone marrow tissue, Environ. Health Perspect. 111 (2003) 1411–1420.

[23] International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome, Nature 431 (2004) 931–945.
[24] R.E. Banks, M.J. Dunn, D.F. Hochstrasser, J.C. Sanchez, W. Blackstock, D.J. Pappin, P.J. Selby, Proteomics: new perspectives, new biomedical opportunities, Lancet 356 (2000) 1749–1756.
[25] N. Tang, P. Tornatore, S.R. Weinberger, Current developments in SELDI affinity technology, Mass Spectrom. Rev. 23 (2004) 34–44.
[26] W.A. Joo, D. Sul, D.Y. Lee, E. Lee, C.W. Kim, Proteomic analysis of plasma proteins of workers exposed to benzene, Mutat. Res. 558 (2004) 35–44.
[27] Z. Wang, C. Yip, Y. Ying, J. Wang, X.Y. Meng, L. Lomas, T.T. Yip, E.T. Fung, Mass spectrometric analysis of protein markers for ovarian cancer, Clin. Chem. 50 (2004) 1939–1942.
[28] K. Gilbert, S. Figueredo, X.Y. Meng, C. Yip, E.T. Fung, Serum protein-expression profiling using the ProteinChip biomarker system, Meth. Mol. Biol. 264 (2004) 259–269.