Chemico-Biological Interactions 184 (2010) 86–93

Contents lists available at ScienceDirect

Chemico-Biological Interactions

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

Review

Systems biology of human benzene exposure

Luoping Zhang[a,*], Cliona M. McHale[a], Nathaniel Rothman[b], Guilan Li[c], Zhiying Ji[a], Roel Vermeulen[d], Alan E. Hubbard[a], Xuefeng Ren[a], Min Shen[b], Stephen M. Rappaport[a], Matthew North[e], Christine F. Skibola[a], Songnian Yin[c], Christopher Vulpe[e], Stephen J. Chanock[b], Martyn T. Smith[a], Qing Lan[b]

[a] Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA 94720-7356, USA
[b] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, 6120 Executive Boulevard, MSC 7242, Bethesda, MD 20892, USA
[c] Chinese Center for Disease Control and Prevention, 27 Nan Wei Road, Beijing 100050, China
[d] Environmental Epidemiology, Institute of Risk Assessment Sciences, Jenalaan 18d, 3584 CK, Utrecht, Netherlands
[e] Department of Nutritional Science and Toxicology, University of California, Berkeley, 119 Morgan Hall #3104, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:
Available online 21 December 2009

Keywords:
Benzene
Toxicogenomics
Epigenetics
Systems biology
Hematotoxicity
Human

ABSTRACT

Toxicogenomic studies, including genome-wide analyses of susceptibility genes (genomics), gene expression (transcriptomics), protein expression (proteomics), and epigenetic modifications (epigenomics), of human populations exposed to benzene are crucial to understanding gene–environment interactions, providing the ability to develop biomarkers of exposure, early effect and susceptibility. Comprehensive analysis of these toxicogenomic and epigenomic profiles by bioinformatics in the context of phenotypic endpoints, comprises systems biology, which has the potential to comprehensively define the mechanisms by which benzene causes leukemia. We have applied this approach to a molecular epidemiology study of workers exposed to benzene. Hematotoxicity, a significant decrease in almost all blood cell counts, was identified as a phenotypic effect of benzene that occurred even below 1 ppm benzene exposure. We found a significant decrease in the formation of progenitor colonies arising from bone marrow stem cells with increasing benzene exposure, showing that progenitor cells are more sensitive to the effects of benzene than mature blood cells, likely leading to the observed hematotoxicity. Analysis of transcriptomics by microarray in the peripheral blood mononuclear cells of exposed workers, identified genes and pathways (apoptosis, immune response, and inflammatory response) altered at high (>10 ppm) and low (<1 ppm) benzene levels. Serum proteomics by SELDI-TOF-MS revealed proteins consistently down-regulated in exposed workers. Preliminary epigenomics data showed effects of benzene on the DNA methylation of specific genes. Genomic screens for candidate genes involved in susceptibility to benzene toxicity are being undertaken in yeast, with subsequent confirmation by RNAi in human cells, to expand upon the findings from candidate gene analyses. Data on these and future biomarkers will be used to populate a large toxicogenomics database, to which we will apply bioinformatic approaches to understand the interactions among benzene toxicity, susceptibility genes, mRNA, and DNA methylation through a systems biology approach.

Contents

1. Introduction to systems biology 87
2. Application of systems biology in studies of benzene toxicity 87

Abbreviations: AML, Acute myeloid leukemia; BFU-E, Burst forming unit-erythroid; BiNGO, Biological network gene ontology; BT, 1,2,4-Benzenetriol; CAT, Catechol; CBC, Complete blood counts; CFU-GEMM, Colony forming unit-granulocyte, erythrocyte, monocyte, megakaryocyte; CFU-GM, Colony forming unit-granulocyte, monocyte; CYP2E1, Cytochrome P450 2E1; GWAS, Genome-wide association studies; HQ, Hydroquinone; IARC, International Agency for Research on Cancer; MDS, Myelodysplastic syndromes; miRNA, MicroRNA; MPO, Myeloperoxidase; NHL, Non-Hodgkin lymphoma; NQO1, NAD(P)H: quinone oxidoreductase 1; PBMC, Peripheral blood mononuclear cells; PF4, Platelet factor 4; q-PCR, Quantitative-polymerase chain reaction; RNAi, RNA interference; SELDI-TOF-MS, Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry; shRNA, Short hairpin RNA; siRNA, Small interfering RNA; SNPs, Single-nucleotide polymorphism; t-MDS/AML, Therapy-related MDS and AML; WBC, White blood cells.

* Corresponding author at: Division of Environmental Health Sciences, School of Public Health, University of California, 237 Hildebrand Hall, Berkeley, CA 94720-7356, USA. Tel.: +1 510 643 5189; fax: +1 510 642 0427.
E-mail address: luoping@berkeley.edu (L. Zhang).

0009-2797/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cbi.2009.12.011

3. Hematotoxicity as a phenotypic outcome of benzene exposure 88
4. Gene expression profiling by transcriptomics 89
5. Proteomic biomarkers of benzene 89
6. Epigenomics in pilot benzene studies 89
6.1. DNA methylation array 89
6.2. miRNA microarrays 90
7. Identification of susceptibility genes by genomics 90
7.1. Genomic screening in yeast 90
7.2. Selection of homologous human genes by bioinformatics 90
7.3. Functional testing of candidate genes by RNAi in human cells 91
7.4. Genotyping results in human population studies 91
8. Systems biology approach in current and future studies 91
Conflict of interest 92
Acknowledgements 92
References 92

1. Introduction to systems biology

Systems biology is a recent and evolving interdisciplinary field that focuses on the systematic study of complex interactions in biological systems [1,2]. Systems biology employs a holistic approach to study all components and interactions in the network of DNA (genes), RNA, proteins and biochemical reactions within a cell or organism. This new field utilizes powerful tools that include toxicogenomics, epigenomics, bioinformatics, and phenomics, classical toxicological or phenotypic endpoints (Fig. 1).

Toxicogenomics combines toxicology with molecular profiling technologies, including genomics (DNA), transcriptomics (mRNA), proteomics (proteins) and metabolomics (chemical metabolites) to elucidate molecular mechanisms involved in chemically induced toxicity. Chemically induced alterations in the transcriptome, proteome, and metabolome are analyzed in the context of the stable, inherited genome, which is assessed by genomics. Toxicogenomic studies of human populations are crucial to understanding gene–environment interactions, and can provide the ability to develop novel biomarkers of exposure (exposome), early effect (responsome), and susceptibility (genome) [3–5]. Epigenomics is the study of epigenetic elements, including DNA methylation (methylomics), non-coding microRNA (miRNAomics) along with small interfering RNA (siRNA) and short hairpin RNA (shRNA) for RNA interference (RNAi), and histone modification. Epigenetic modifications play an essential role in regulating gene expression and biological and molecular functions in living cells, without altering the genome.

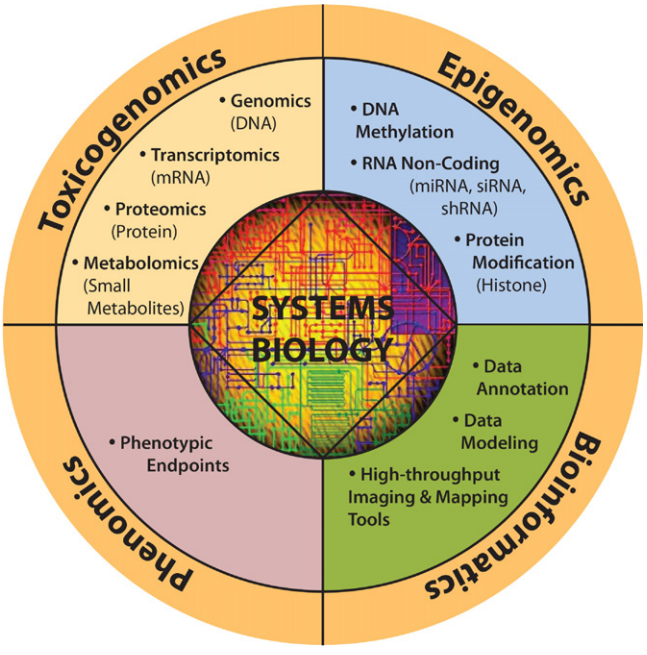

Fig. 1. Overview of systems biology and its components (center of image, Wired Systems Biology, adapted from Chemical & Engineering News 81 (20) 2003).

Another tool central to systems biology is bioinformatics, the application of computational information technology to the field of molecular biology to understand how cells and cell systems work [6,7]. Bioinformatics facilitates the analysis of complex biological data (toxicogenomic and epigenomic endpoints) and applies knowledge from annotated functions, pathways and networks to describe normal and perturbed biological states, also known as phenomics, the study of outcomes (phenotypic endpoints). Together, these omic technologies can each provide a "molecular signature" or "fingerprint" of chemical exposure, early effect or genetic susceptibility, which may enhance our understanding of gene–environment interactions. Thus, this holistic approach known as systems biology has the potential to comprehensively define the mechanisms contributing to disease. The purpose of this review paper is to describe how and why it is important to apply the "Systems Biology" approach to benzene mechanistic studies and future directions.

2. Application of systems biology in studies of benzene toxicity

Benzene, a ubiquitous chemical, is an established cause of acute myeloid leukemia (AML), myelodysplastic syndromes (MDS), and probably lymphocytic leukemias and non-Hodgkin lymphoma (NHL) in humans [8–11]. Benzene-induced toxicity in blood-forming systems has been known for more than a century [12]. In 1982, the International Agency for Research on Cancer (IARC) stated, "*There is sufficient evidence that benzene is carcinogenic to man*" [13], and when a new IARC classification system was established in 1987, benzene was immediately placed in the Group 1 human carcinogen category [14].

Potential mechanisms of benzene toxicity have been investigated primarily in the following areas [15]: (1) benzene metabolism in the liver (CYP2E1, etc.) and transportation to the bone marrow for secondary metabolism (MPO, NQO1) [16,17]; (2) oxidative stress from reactive oxygen species generated by redox cycling [18,19]; (3) chromosome alterations including translocations, deletions, and aneuploidy [12]; (4) protein damage to tubulin, histone proteins, topoisomerase II, etc. [15]; and (5) immune system dysfunction (TNF-α, INF-γ, AhR, etc.) [20–22]. Benzene induces chromosomal alterations similar to those found in therapy-related MDS and AML (t-MDS/AML), and in *de novo* leukemia [23]. Distinct chromosome effects arise following exposure to alkylating

agents (5q-/-5 or 7q-/-7 and associated genetic abnormalities) and topoisomerase II inhibitors (recurrent balanced translocations or inversions) used in chemotherapy treatment. Exposure to benzene or its metabolites has been associated with loss and long (q) arm deletion of chromosomes 5 and 7 [24] and translocations involving t(21q) [25,26], further suggesting that benzene induces leukemia through multiple different mechanisms.

Studies to date have provided evidence for multiple potential mechanisms using approaches that rely on limited research tools that analyze only one or a few, *a priori* selected genes, pathways or metabolites at a time. A systems biology approach is necessary to interrogate all potential mechanisms by which benzene exposure contributes to disease, through the application of unbiased omic-based technologies in an integrated manner. Since the last international benzene conference at Munich in 2004, we have adopted such an approach to understand the mechanisms underlying human benzene toxicity. This review summarizes our findings published over the last 5 years and preliminary data from recent pilot studies (Table 1). It also provides an overview of our current understanding of benzene-induced hematotoxicity and suggestions for further research.

We describe the studies in the context of systems biology as defined in Fig. 1. First, we discuss hematotoxicity as a phenotypic outcome of benzene exposure, with implications for adverse future health effects. We then discuss findings from two toxicogenomic studies, transcriptomics and proteomics, followed by preliminary epigenomics data. As these toxicogemomic and epigenomic responses to exposure are likely influenced by susceptibility, next we describe how we investigated human susceptibility genes using a yeast genomic screening approach with validation of homologous human genes in human cells. We also describe genotyping studies of candidate genes in human exposed populations. Finally, we discuss how the current and future omic datasets could be integrated, using sophisticated bioinformatics approaches in progress, into one consolidated model of the perturbations effected by benzene. This could identify robust biomarkers and help to clarify the molecular and cellular networks impacted by benzene, yielding a more comprehensive understanding of the mechanistic effects of benzene.

3. Hematotoxicity as a phenotypic outcome of benzene exposure

Although benzene was known to have toxic effects on the hematopoietic system (hematotoxicity) at high, occupational doses for over a century [12], the degree of hematotoxicity at low levels of exposure was largely unknown. Recently, a study of 250 workers exposed to varying levels of benzene and 140 unexposed controls in Tianjin, China, during which benzene and other chemical exposure levels were monitored repeatedly for up to 12 months, was conducted. Air, urine and blood samples were collected and complete blood counts (CBC) analyzed [27]. In comparison with the non-exposed controls ($n = 140$), a significant decrease was observed in almost all blood cell counts, such as white blood cells (WBC), granulocytes, lymphocytes and platelets, in workers exposed to benzene ($n = 250$), even at exposures below 1 ppm ($n = 109$), the current occupational standard in the U.S. Additionally, lymphocyte subset analysis showed significant, dose-dependent, decreases in $CD4^+$-T cells, $CD4^+/CD8^+$ ratio, and B cells at <1 ppm benzene exposures. These findings, based on the differentiated blood cell counts, provide evidence of bone marrow toxicity in workers exposed to benzene at or below 1 ppm [27].

Because all types of WBC counts were suppressed, it was suspected that the number or functionality of hematopoietic stem and/or progenitor cells generated in the bone marrow had been

Table 1
Summary of benzene studies applying systems biology approach.

Systems biology	Study method	Study size (n)		Benzene exposure (ppm)	Major findings	Reference
		Exposed	Controls			
Phenomics						
Hematotoxicity	CBC	250	140	<1, <10, >10	Decrease in all blood cell counts	[27]
	Colony	24	29	<10, >10	Decrease in colony formation	[27]
Toxicogenomics						
Transcriptomics	Affymetrix	6	6	≥10	29 genes differentially expressed	[29]
	q-PCR	13	15	≥10	Validated 6 of 29 genes	
	Affymetrix and Illumina	8	8		Confirms Forrest findings	[30]
	Illumina	83	42	<1, <10, >10	Different gene pathways identified at low and high exposures	[31]
Proteomics	SELDI-TOF	20 (2 sets of 10)	20	31.3 and 37.9 (mean)	Identified two down-regulated proteins as PF4 and CTAP-III	[32]
Epigenomics						
DNA methylation	Illumina (GoldenGate)	6	4	8.9 ± 9.1 (mean ± SD)	Results preliminary (*MSH3*, *RUNX3*)	Pilot study
		Human TK6 cells		HQ (0, 10, 15, 20 μM)	Results preliminary (*RUNX1*, *IL12*)	
miRNA	Agilent	7	7	<1 ppm	Results preliminary (four miRNAs)	Pilot study
Genomics	PDA in yeast	~4600 homozygous deletion strains		HQ, CAT, BT	Oxidative stress response	[44]
	RNAi in human cells	WRN		HQ	DNA repair HR pathway	[48,49]
Genotyping	Illumina (GoldenGate)	250	140	<1, <10, >10	Identified *MPO*, *NQO1*, and a group of DNA repair and cytokine genes	[50–52]

reduced by benzene. To test this hypothesis, we cultured the progenitor cells circulating in peripheral blood and examined the effects of benzene on different types of progenitor cell colony formation (CFU-GM, BFU-E, CFU-GEMM). The results showed highly significant dose-dependent decreases in colony formation from all three types of progenitor cells, especially when compared to the corresponding decreased levels of differentiated WBC and granulocytes [27]. This suggests that early myeloid progenitor cells are more sensitive than mature cells to the effects of benzene, and clarifies the role of benzene in reduced blood cell counts.

Overall, these hematologic effects could reflect events in the bone marrow that may be associated with adverse health effects in the future [28]. Having established that hematotoxicity, specifically effects myeloid progenitor cells, as a phenotypic anchor of benzene toxicity, we began to examine the molecular mechanisms underlying these effects, through the comprehensive systems biology approach proposed above.

4. Gene expression profiling by transcriptomics

Transcriptomic studies are useful in determining the impact of environmental or occupational exposure to chemicals on the transcriptome, the set of all mRNA transcripts expressed within a cell. To better understand the risks of benzene in humans, the peripheral blood mononuclear cell (PBMC) transcriptomes from occupationally exposed workers in China were examined by microarray (Affymetrix). Analysis of six exposed-control pairs revealed differential expression in 29 genes in the exposed individuals, compared to the controls. Four genes, *CXCL16*, *ZNF331*, *JUN*, and *PF4*, were shown to be potential biomarkers of early response to benzene exposure as they were confirmed by quantitative-polymerase chain reaction (q-PCR) [29].

A later study of eight exposed-control pairs confirmed these results, using two different microarray platforms (Affymetrix and Illumina) to identify global gene expression changes. The differential expression of 2692 genes and 1828 genes was found by Affymetrix and Illumina, respectively, and the 4 genes, *CXCL16*, *ZNF331*, *JUN*, and *PF4*, were among the most significantly altered, validating the findings from the earlier Forrest et al. study. This study additionally identified biological pathways that were associated with high benzene exposure, including genes involved in apoptosis and lipid metabolism. This study used a two-platform approach that identified robust changes in the PBMC transcriptome of benzene-exposed individuals [30].

The effects of exposure to high levels of benzene are well documented compared to low-level exposure, the latter being more challenging due to confounders. More recently, we have shown, in an expanded study of 125 factory workers, that low-dose benzene exposure (<1 ppm, $n = 59$) can also cause widespread subtle, yet highly significant, perturbation of gene expression in PBMC. This study was designed with sufficient power to detect robust expression changes, accounting for technical variability as well as age, gender and other confounders. Our microarray analysis revealed significant dysregulation of more than 2500 genes by low-dose benzene exposure, over 70 of which had differential expression ratios exceeding 1.5. Several of the detected genes exhibited significantly altered expression only at low levels of benzene exposure, and are thus potential biomarkers of low-dose exposure. The findings show that even low levels of occupational benzene exposure cause a significant perturbation of expression of genes involved in immune and inflammatory responses [31].

5. Proteomic biomarkers of benzene

Another important toxicogenomic tool, proteomics, can be used to measure alterations in the proteome (e.g. protein levels, post-translational modifications) associated with exogenous chemical exposure. Effects on the blood proteome may reflect effects at distal body sites. As with transcriptomics, proteomics can be used to discover biomarkers of exposure and early effect, as well as increase our understanding of the mechanisms underlying disease.

We examined the impact of benzene on the human serum proteome in exposed factory workers and controls to obtain insight into the mechanism of action of benzene [32]. Serum samples were fractionated and proteins were bound to surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) chips. Three proteins (4.1, 7.7, and 9.3 kDa) were consistently down-regulated in the exposed ($n = 10$) compared to the control ($n = 10$) individuals in two separate sets of study subjects (40 subjects total). All proteins were highly inversely correlated with individual estimates of benzene exposure. The 7.7- and 9.3-kDa proteins were identified as platelet factor 4 (PF4), also down-regulated at the gene expression level described above, and connective tissue activating peptide CTAP-III, respectively, both platelet-derived CXC chemokines. Thus, reduced protein levels of PF4 or CTAP-III are potential biomarkers of the early biologic effects of benzene. Future proteomic studies could identify further biomarkers of benzene exposure, and elucidate the mechanisms underlying benzene toxicity and associated disease.

6. Epigenomics in pilot benzene studies

Gene expression and ultimately protein expression is regulated at the epigenetic level by processes including DNA methylation, histone modification and miRNA (microRNA) expression. The epigenome, while stable through cell division and even in some cases reproduction, can be reprogrammed by nutritional, chemical, and physical factors [33]. Thus, the study of toxic effects on the epigenome is crucial to understanding mechanisms of action. Further, epigenetic modifications represent more stable biomarkers and fingerprints of exposure than altered gene or protein expression [34]. While epigenetics refers to the study of individual or specific gene activity, epigenomics focuses on global analyses of epigenetic changes across the entire genome. A recent study reported that hyper-methylation in *p15* and genome-wide hypo-methylation assessed by LINE-1 (Long Interspersed Nuclear Element-1) were associated with very low benzene exposures (~22 ppb), in healthy subjects including gas station attendants and traffic police officers, although the corresponding effects on methylation were very low [35]. To determine whether epigenetics plays a role in the hematotoxicity of benzene, we have recently performed several pilot epigenomic studies including DNA methylation and miRNA expression arrays in the blood of workers occupationally exposed to benzene. The results described below are very preliminary, and serve mainly as a proof of principle of this epigenomic approach.

6.1. DNA methylation array

A DNA methylation array (GoldenGate Methylation Cancer Panel I, Illumina) was applied to determine the methylation status of the CpG islands of >800 genes in DNA isolated from the buffy coats of 6 benzene-exposed workers (2 male, 4 female) and 4 unexposed controls (2 male, 2 female). As expected gender-specific methylation patterns were seen for numerous genes including *ELK1*, *EFNB1*, *MYCL2*, *VBP1*, *DNASE1L1*, *DKC1* and *CDM*. This pilot study also found altered methylation induced by benzene at many CpG sites. Decreased methylation of *RUNX3* (*AML2*), a gene whose altered expression has been associated with myeloproliferative disorders [36] occurred at three different CpG sites (Fig. 2A). Increased methylation of *MSH3*, a critical gene in the maintenance of genome

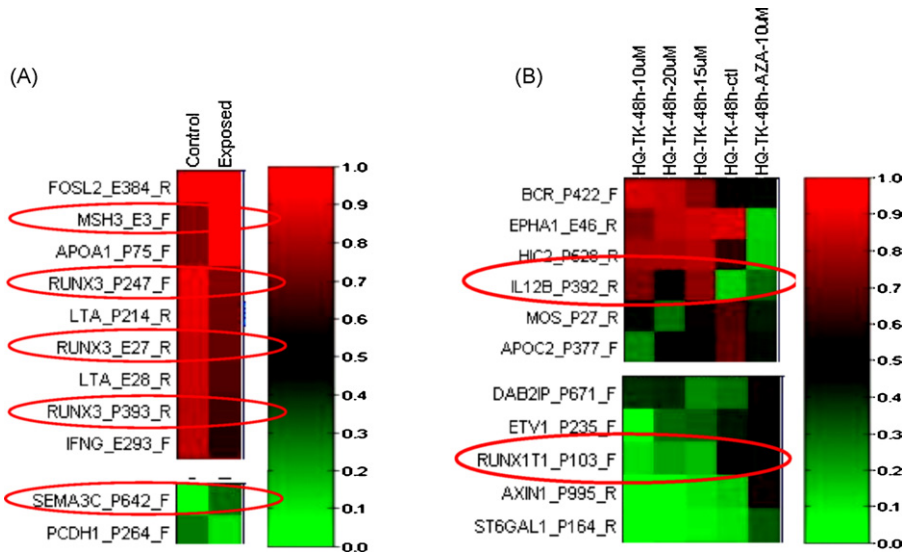

Fig. 2. Heatmap of methylation profiles of (A) workers exposed to benzene and controls, and (B) TK6 cells exposed to hydroquinone (HQ). (A) A total of six benzene-exposed subjects and four controls were analyzed. (B) Human TK6 cells were treated with HQ at 0, 10, 15 and 20 μM for 48 h. 5-Azacytidine, a demethylating agent, was included as a positive control. Examples of genes with methylation levels significantly altered by benzene or HQ are shown. Values range from 1.0 fully methylated (red) to 0 fully unmethylated (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

integrity, and *Sema3C*, a secreted guidance protein implicated in tumorigenesis [37], was also observed (Fig. 2A). There appeared to be a gender-specific effect of benzene on the methylation of several genes, although the sample number was small.

We examined the methylation status of the same panel of genes in DNA from TK6 cells treated *in vitro* with hydroquinone (HQ) at 10, 15, and 20 μM, for 48 h. The most significantly hyper-methylated gene was *IL12* (Fig. 2B), whose expression has previously been shown to be down-regulated by HQ in mouse macrophages [38]. *RUNX1T1* (runt-related transcription factor 1, also known as *ETO*), was hypo-methylated by HQ *in vitro* (Fig. 2B). The protein encoded by this gene is commonly fused to *RUNX1* (*AML1*) in the t(8;21)(q22;q22) translocation, which is one of the most frequent karyotypic abnormalities in AML [39]. *MAGEA1* (melanoma antigen family A, 1), also known as *MAGE-1*, was also hypo-methylated by HQ in the present study (data not shown). Interestingly, this gene was reported recently to be hypo-methylated weakly in subjects exposed to increasing airborne benzene levels [35].

6.2. miRNA microarrays

Human miRNA microarrays (Agilent), containing probes for 470 human and 64 human viral miRNAs, were used to analyze the differential expression of miRNAs in the total PBMC RNA from seven exposed-control matched pairs, in a pilot study. Preliminary analysis showed upregulation of four miRNAs (miR-154*; miR-487a; miR-493-3p; and, miR-668) by benzene exposure. Upregulation of miR-154* expression, possibly through a change in the methylation and acetylation status of the 14q32 region, has been reported in patients with acute promyelocytic leukemia bearing the t(15;17) translocation [40].

While both of these studies are relatively small, and the data are very preliminary, the findings suggest that further studies to examine the epigenetic effects of benzene on gene-specific promoter methylation and miRNA expression, in a larger study of exposed workers, are warranted.

7. Identification of susceptibility genes by genomics

Genome-wide association studies (GWAS) are a well-established means to examine the association of genetic susceptibility, i.e. single-nucleotide polymorphisms (SNPs), with disease. Such studies have been performed for two diseases related to benzene toxicity, NHL [41] and t-AML [42]. A very small number of pharmacogenomic GWAS have been reported [43], while to our knowledge, no GWAS on occupational or environmental exposure and associated toxic outcomes have been performed. Given the relatively small effects observed in disease GWAS, such studies are very expensive to undertake for less well-defined, pre-disease, toxicological outcomes. We adopted an alternative genomic approach to discover human susceptibility genes, the aims of which are: (1) to discover/screen susceptible genes in yeast by genomics; (2) to select human homolog genes using bioinformatics; (3) to test identified gene (e.g. *WRN*) functions by RNAi in human cells; 4) to identify human susceptible genes by SNP genotyping in population studies.

7.1. Genomic screening in yeast

In order to reduce the complexity and expense of analyzing the human genome, while retaining the ability to systematically screen a genome highly relevant to human biology, we chose a screening system in yeast (*Saccharomyces cerevisiae*). As we reported recently, this genomics approach has been employed to discover novel biomarkers of benzene toxicity in yeast cells exposed to the major active metabolites of benzene, hydroquinone (HQ), catechol (CAT), and 1,2,4-benzenetriol (BT) [44]. Using a collection of yeast strains representing a complete set of non-essential gene deletions, genetically tagged so that individual strains can be identified in competitive growth experiments, fitness assays were performed to identify mutant strains whose fitness is significantly altered following treatment with benzene metabolites. A comparison of the global deletome profiles of the metabolites revealed that deletion of certain genes rendered yeast cells sensitive to all three compounds. Several of the genes identified in the yeast studies have human orthologs with conserved biological function, supporting the notion that the mechanisms of toxicity identified in yeast are relevant to human disease.

7.2. Selection of homologous human genes by bioinformatics

To select and prioritize likely human candidate genes from complex yeast genomic data, we applied bioinformatic analyses

using specific computational programs including "clustering" by HOPACH (Hierarchical Ordered Partitioning And Collapsing Hybrid) algorithm methods [45], pathway analysis using Cytoscape with the BiNGO Gene Ontology identification plugin [46,47] and a comparative genomics approach. It is suggested that toxicants of similar mechanisms of action most likely have similar profiles of genes required for tolerance. We thus employed a variety of computational "clustering" methods to analyze the sensitivity and resistance data and to identify yeast strains most sensitive to each of the benzene metabolites tested as well as the genes specifically involved in sensitivity to each toxicant. We also set out to identify biologically significant patterns and features involved in toxicant response between metabolites. Similarly, a comparative genomics approach has been applied to identify functional orthologs and pathways between evolutionarily distant organisms. This approach assists in the identification of yeast and subsequently human candidate genes and pathways for further evaluation in the mammalian cell culture system.

7.3. Functional testing of candidate genes by RNAi in human cells

The roles of the human homologs of selected genes in benzene toxicity have been examined through mechanistic studies in human cell lines. SNPs in WRN, an important protein that plays a role in the maintenance of genomic stability, have been associated with an increased risk for some cancers and benzene hematotoxicity. We knocked down WRN protein using siRNA in HeLa cells and examined sensitivity to toxicity following exposure to the benzene metabolite, HQ [48]. Depletion of WRN led to decreased cell proliferation and increased HQ cytotoxicity, evident by increased necrosis. Additionally, these cells displayed increased DNA double-strand breaks (DSB), a potential biomarker of benzene hematotoxicity. Together, the results showed that WRN plays an important role in resistance to benzene toxicity in HeLa, and perhaps other cells.

More recently, we used shRNA to silence WRN in the human HL60 acute promyelocytic cell line [49]. Upon exposure to HQ, HL60 cell growth rates were accelerated, and DNA breaks and sensitivity to HQ-induced cytotoxicity and genotoxicity were increased, similar to the findings in HeLa cells. Loss of WRN also resulted in higher levels of early apoptosis. An accumulation of such genetic lesions can lead to the development of AML. The data from this study provides mechanistic support for the link between WRN and benzene-induced hematotoxicity, and possibly even benzene-induced leukemia. Studies are underway to confirm the role of other susceptibility genes in benzene toxicity.

7.4. Genotyping results in human population studies

In addition to the yeast genomic studies, large-scale human population genotyping studies have also been conducted. In collaboration between the National Cancer Institute and China CDC, such a study analyzed 1395 SNPs in 411 potential carcinogenesis-related genes using an Illumina GoldenGate assay in 250 benzene-exposed workers and 140 unexposed controls in China [50]. One or more SNPs in five genes (*WRN*, *BLM*, *TP53*, *RAD51*, and *WDR79*) which play critical roles in DNA repair and genomic maintenance, were associated with highly significant 10–20% reductions (p values ranged from 0.0011 to 0.0002) in the WBC count among benzene-exposed workers but not controls, with evidence for gene–environment interactions for SNPs in *BLM*, *WRN* and *RAD51*. Earlier candidate gene studies identified a small number of SNPs in genes involved in benzene metabolism, cytokine and cellular adhesion molecules [51], and DNA DBS repair [52], which appear to confer susceptibility to benzene hematotoxicity. These studies were the first to provide evidence that genetic polymorphisms in certain genomic stability maintenance genes, like *WRN* [53], impact benzene-induced phenotypic outcomes such as hematotoxicity.

The human population studies confirm a critical role for DNA repair and genomic maintenance in susceptibility, and further support these effects as benzene-induced phenotypic outcomes. In addition, genetic variants in metabolizing enzymes responsible for activating and detoxifying benzene, in particular *MPO* and *NQO1*, have also been linked to increased susceptibility to benzene hematotoxicity [27]. Together, these genomic and genotyping studies provide important information regarding benzene toxicity and disease pathways.

8. Systems biology approach in current and future studies

From our omic studies to date, benzene appears to cause hematotoxicity through multiple mechanisms that may involve alterations in the expression of multiple genes and proteins, DNA methylation patterns and miRNA profiles even at low-doses. Transcriptomics has identified many genes, functions and pathways altered by benzene, offering insight into mechanisms and providing potential signatures of benzene exposure, and/or early effect. These data could be integrated with information on susceptibility genes to further understand gene–environment interactions and perhaps to identify the most susceptible individuals. We will expand our proteomic studies by conducting further analyses on different protein fractions and affinity chips to identify more altered proteins. As discussed earlier, expanded DNA methylation and miRNA profiling studies are necessary in larger populations of exposed individuals. All studies will be performed at a range of benzene exposures to examine dose–response effects. Several of these individual omic datasets are large, measuring, e.g. the expression level of ~24,000 genes or the methylation status of >14,000 genes at multiple CpG islands and bioinformatic methods to analyze them continue to be refined.

Each individual omic dataset is anticipated to provide information on the effects of benzene, and potentially identify biomarkers of exposure and early effect. Through our systems biology approach, we will use sophisticated bioinformatics to integrate individual datasets into one consolidated model of the perturbations effected by benzene. From this model we will make inferences, specifically, we will aim to understand the interactions between benzene toxicity, SNPs, mRNA, miRNA, protein, and DNA methylation. This could identify robust biomarkers and help to clarify the molecular and cellular networks impacted by benzene, yielding a more comprehensive understanding of the mechanistic effects of benzene. While all the toxicogenomic endpoints have the potential to yield biomarkers, some endpoints such as DNA methylation and gene expression may reflect more upstream mechanistic effects while others such as proteomics may reflect more downstream, phenotypic effects, and might be more informative of actual molecular and cellular processes affected. Multiple regulatory mechanisms probably determine the phenotypic outcome (e.g. a gene could be up-regulated by DNA methylation and down-regulated by miRNA). The systems biology approach will require a high level of computing power and will capitalize on the ever-expanding knowledge of biological pathways and networks. In order to realize this approach, future studies need to be designed with sufficient power to robustly detect effects of benzene and to allow for analysis of the inter-relationship among the different endpoints. With respect to sample processing, the ideal scenario is to analyze DNA, RNA and protein from the same cell population but this remains challenging with existing protocols and sample availability.

The findings from the systems biology study of benzene could also contribute more generally to the field of risk assessment. Comparison of the toxicogenomic, epigenomic and genomic profiles associated with different exposures, e.g. suspected leukemogens or carcinogens, and diseases (NHL, AML), may help to clarify the connection between chemicals, genes/proteins, pathways/networks, and disease. Initiatives such as the Comparative Toxicogenomics Database [54] and Chemical Effects in Biological Systems [55] have been developed towards this goal. It has been also discussed how omic data/measurements obtained through a systems biology approach can be applied to identify all potential mechanisms of action and serve as an information base for subsequent evaluation of these mechanisms when conducting risk assessment [56].

In conclusion, the systems biology approach described here should help inform the mechanisms underlying benzene hematoxicity and associated disease, and identify robust biomarkers of exposure, early effect, susceptibility and disease development.

Conflict of interest

MTS and SMR have received consulting and expert testimony fees from law firms in cases involving exposure to benzene. GL has been a consultant for the American Petroleum Institute on benzene-related health research.

Acknowledgements

We thank the participants for taking part in this study. We are extremely grateful to Heather Lustig and Kshama Mehta at Agilent Technologies, for facilitating the miRNA pilot study. We thank Ms. Suzanne May for her technical assistance with the DNA methylation analysis. We are grateful to Ms. Anh Duong for her tireless contribution and editorial assistance to this manuscript, and to Mr. Patrick Chang for his technical talent to illustrate the Systems Biology figure. This work was supported by NIH grants RO1ES06721 and P42ES04705 (MTS), P42ES05948 and P30ES10126 (SMR).

References

[1] M.W. Kirschner, The meaning of systems biology, Cell 121 (2005) 503–504.
[2] H. Kitano, Systems biology: a brief overview, Science (New York, NY) 295 (2002) 1662–1664.
[3] C. McHale, L. Zhang, A.E. Hubbard, M. Smith, Toxicogenomic studies in human populations, in: D.R. Boverhof, B.B. Gollapudi (Eds.), Applications of Toxicogenomics in Safety Evaluation and Risk Assessment, John Wiley & Sons, Inc., Hoboken, NJ, 2010.
[4] M. Smith, S. Rappaport, Building exposure biology centers to put the E into "G × E" interaction studies, Environ. Health Perspect. 117 (2009) A334–A335.
[5] J. Vlaanderen, L.E. Moore, M.T. Smith, et al., Application of omics technologies in occuptional and environmental health research: current status and projections, Occup. Environ. Med. (2009), http://www.ncbi.nlm.nih.gov/pubmed/19933307?itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum&ordinalpos=2.
[6] A. Bayat, Science, medicine, and the future: bioinformatics, BMJ 324 (2002) 1018–1022.
[7] National Research Council (NRC), Toxicity Testing in the 21st Century: A Vision and a Strategy, Committee on Toxicity Testing and Assessment of Environmental Agents, National Academies Press, Washington, DC, 2007.
[8] D.C. Glass, C.N. Gray, D.J. Jolley, et al., Leukemia risk associated with low-level benzene exposure, Epidemiology (Cambridge, MA) 14 (2003) 569–577.
[9] R.B. Hayes, Y. Songnian, M. Dosemeci, M. Linet, Benzene and lymphohematopoietic malignancies in humans, Am. J. Ind. Med. 40 (2001) 117–126.
[10] D.A. Savitz, K.W. Andrews, Review of epidemiologic evidence on benzene and lymphatic and hematopoietic cancers, Am. J. Ind. Med. 31 (1997) 287–295.
[11] M.T. Smith, R.M. Jones, A.H. Smith, Benzene exposure and risk of non-Hodgkin lymphoma, Cancer Epidemiol. Biomark. Prev. 16 (2007) 385–391.
[12] L. Zhang, D.A. Eastmond, M.T. Smith, The nature of chromosomal aberrations detected in humans exposed to benzene, Crit. Rev. Toxicol. 32 (2002) 1–42.
[13] International Agency for Research on Cancer (IARC), Some Industrial Chemicals and Dyestuffs, Monogr. Eval. Carcinog. Risks Hum., 1982.
[14] International Agency for Research on Cancer (IARC), Monographs on the Evaluation of Carcinogenic Risks to Humans Overall Evaluations of Carcinogenicity, An Updating of IARC Monographs, 1987.
[15] M.T. Smith, The mechanism of benzene-induced leukemia: a hypothesis and speculations on the causes of leukemia, Environ. Health Perspect. 104 (Suppl. 6) (1996) 1219–1225.
[16] D. Ross, D. Siegel, D.G. Schattenberg, X.M. Sun, J.L. Moran, Cell-specific activation and detoxification of benzene metabolites in mouse and human bone marrow: identification of target cells and a potential role for modulation of apoptosis in benzene toxicity, Environ. Health Perspect. 104 (Suppl. 6) (1996) 1177–1182.
[17] R. Snyder, C.C. Hedli, An overview of benzene metabolism, Environ. Health Perspect. 104 (Suppl. 6) (1996) 1165–1171.
[18] V.V. Subrahmanyam, D. Ross, D.A. Eastmond, M.T. Smith, Potential role of free radicals in benzene-induced myelotoxicity and leukemia, Free Radic. Biol. Med. 11 (1991) 495–515.
[19] L. Zhang, M. Smith, B. Bandy, S.J. Tamaki, A.J. Davison, Role of Quinones, active oxygen species and metals in the genotoxicity of 1,2,4-benzenetriol, a metabolite of benzene, in: H. Nhol, H. Esterbauer, C. Rice-Evans (Eds.), Free Radicals in the Environment, Medicine and Toxicology, Richelieu, London, 1994, pp. 521–562.
[20] J.Y. Cho, Suppressive effect of hydroquinone, a benzene metabolite, on in vitro inflammatory responses mediated by macrophages, monocytes and lymphocytes, Mediat. Inflamm. (2008) 298010.
[21] Y. Ouyang, N. Virasch, P. Hao, et al., Suppression of human IL-1beta, IL-2, IFN-gamma, and TNF-alpha production by cigarette smoke extracts, J. Allergy Clin. Immunol. 106 (2000) 280–287.
[22] B.I. Yoon, Y. Hirabayashi, Y. Kawasaki, et al., Aryl hydrocarbon receptor mediates benzene-induced hematotoxicity, Toxicol. Sci. 70 (2002) 150–156.
[23] J. Pedersen-Bjergaard, M.K. Andersen, M.T. Andersen, D.H. Christiansen, Genetics of therapy-related myelodysplasia and acute myeloid leukemia, Leukemia 22 (2008) 240–248.
[24] L. Zhang, N. Rothman, Y. Wang, et al., Increased aneusomy and long arm deletion of chromosomes 5 and 7 in the lymphocytes of Chinese workers exposed to benzene, Carcinogenesis 19 (1998) 1955–1961.
[25] Z. Ji, L. Zhang, W. Guo, C.M. McHale, M.T. Smith, The benzene metabolite, hydroquinone and etoposide both induce endoreduplication in human lymphoblastoid TK6 cells, Mutagenesis 24 (2009) 367–372.
[26] M.T. Smith, L. Zhang, Y. Wang, et al., Increased translocations and aneusomy in chromosomes 8 and 21 among workers exposed to benzene, Cancer Res. 58 (1998) 2176–2181.
[27] Q. Lan, L. Zhang, G. Li, et al., Hematotoxicity in workers exposed to low levels of benzene, Science (New York, NY) 306 (2004) 1774–1776.
[28] Q. Lan, R. Vermeulen, L. Zhang, et al., Benzene exposure and hematotoxicity, Science (New York, NY) 312 (2006) 998–999.
[29] M.S. Forrest, Q. Lan, A.E. Hubbard, et al., Discovery of novel biomarkers by microarray analysis of peripheral blood mononuclear cell gene expression in benzene-exposed workers, Environ. Health Perspect. 113 (2005) 801–807.
[30] C.M. McHale, L. Zhang, Q. Lan, et al., Changes in the peripheral blood transcriptome associated with occupational benzene exposure identified by cross-comparison on two microarray platforms, Genomics 93 (2009) 343–349.
[31] C. McHale, L. Zhang, Q. Lan, et al., Low-dose, occupational exposure to the leukemogen benzene induces robust changes in the blood transcriptome associated with altered immune system biology, American Society of Hematology, San Francisco, CA, 2008.
[32] R. Vermeulen, Q. Lan, L. Zhang, et al., Decreased levels of CXC-chemokines in serum of benzene-exposed workers identified by array-based proteomics, Proc. Natl. Acad. Sci. U.S.A. 102 (2005) 17041–17046.
[33] D.C. Dolinoy, R.L. Jirtle, Environmental epigenomics in human health and disease, Environ. Mol. Mutagen. 49 (2008) 4–8.
[34] R.L. Jirtle, M.K. Skinner, Environmental epigenomics and disease susceptibility, Nat. Rev. 8 (2007) 253–262.
[35] V. Bollati, A. Baccarelli, L. Hou, et al., Changes in DNA methylation patterns in subjects exposed to low-dose benzene, Cancer Res. 67 (2007) 876–880.
[36] C.I. Muller, Q.T. Luong, L.Y. Shih, et al., Identification of marker genes including RUNX3 (AML2) that discriminate between different myeloproliferative neoplasms and normal individuals, Leukemia 22 (2008) 1773–1778.
[37] G. Neufeld, N. Shraga-Heled, T. Lange, et al., Semaphorins in cancer, Front. Biosci. 10 (2005) 751–760.
[38] E. Kim, B.Y. Kang, T.S. Kim, Inhibition of interleukin-12 production in mouse macrophages by hydroquinone, a reactive metabolite of benzene, via suppression of nuclear factor-kappaB binding activity, Immunol. Lett. 99 (2005) 24–29.
[39] P Paschka, Core binding factor acute myeloid leukemia, Semin. Oncol. 35 (2008) 410–417.
[40] A. Dixon-McIver, P. East, C.A. Mein, et al., Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia, PLoS ONE 3 (2008) e2141.
[41] C.F. Skibola, P.M. Bracci, E. Halperin, et al., Genetic variants at 6p21.33 are associated with susceptibility to follicular lymphoma, Nat. Genet. 41 (2009) 873–875.
[42] J.A. Knight, A.D. Skol, A. Shinde, et al., Genome-wide association study to identify novel loci associated with therapy-related myeloid leukemia susceptibility, Blood 113 (2009) 5575–5582.
[43] J.J. Crowley, P.F. Sullivan, H.L. McLeod, Pharmacogenomic genome-wide association studies: lessons learned thus far, Pharmacogenomics 10 (2009) 161–163.
[44] M. North, J. Shuga, A. Loguinov, et al., Utilizing functional genomics in yeast to discover novel biomarkers of benzene toxicity in humans, in: International

Symposium on Benzene Health Effects and Mechanisms of Bone Marrow Toxicity, Munich, Germany, 2009.
[45] M.J. Van der Laan, K.S. Pollard, A new algorithm for hybrid hierarchial clustering with visualization and the bootstrap, J. Stat. Plann. Inference (2003) 275–303.
[46] M.S. Cline, M. Smoot, E. Cerami, et al., Integration of biological networks and gene expression data using Cytoscape, Nat. Protoc. 2 (2007) 2366–2382.
[47] S. Maere, K. Heymans, M. Kuiper, BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks, Bioinformatics (Oxford, England) 21 (2005) 3448–3449.
[48] N. Galvan, S. Lim, S. Zmugg, M.T. Smith, L. Zhang, Depletion of WRN enhances DNA damage in HeLa cells exposed to the benzene metabolite, hydroquinone, Mutat. Res. 649 (2008) 54–61.
[49] X Ren, S. Lim, M.T. Smith, L. Zhang, Werner syndrome protein, WRN, protects cells from DNA damage induced by the benzene metabolite hydroquinone, Toxicol. Sci. 107 (2009) 367–375.
[50] Q. Lan, L. Zhang, M. Shen, et al., Large-scale evaluation of candidate genes identifies associations between DNA repair and genomic maintenance and development of benzene hematotoxicity, Carcinogenesis 30 (2009) 50–58.
[51] Q. Lan, L. Zhang, M. Shen, et al., Polymorphisms in cytokine and cellular adhesion molecule genes and susceptibility to hematotoxicity among workers exposed to benzene, Cancer Res. 65 (2005) 9574–9581.
[52] M. Shen, Q. Lan, L. Zhang, et al., Polymorphisms in genes involved in DNA double-strand break repair pathway and susceptibility to benzene-induced hematotoxicity, Carcinogenesis 27 (2006) 2083–2089.
[53] D.K. Singh, B. Ahn, V.A. Bohr, Roles of RECQ helicases in recombination based DNA repair, genomic stability and aging, Biogerontology 10 (2009) 235–252.
[54] C.J Mattingly, M.C. Rosenstein, G.T. Colby, J.N. Forrest Jr., J.L. Boyer, The Comparative Toxicogenomics Database (CTD): a resource for comparative toxicological studies, J. Exp. Zool. A: Comp. Exp. Biol. 305 (2006) 689–692.
[55] M. Waters, S. Stasiewicz, B.A. Merrick, et al., CEBS—Chemical Effects in Biological Systems: a public data repository integrating study design and toxicity data with microarray and proteomics data, Nucleic Acids Res. 36 (2008) D892–900.
[56] S.W. Edwards, R.J. Preston, Systems biology and mode of action based risk assessment, Toxicol. Sci. 106 (2008) 312–318.