EPA/600/R-09/028F | September 2009 | www.epa.gov

United States
Environmental Protection
Agency

An Approach to Using Toxicogenomic Data in U.S. EPA Human Health Risk Assessments: A Dibutyl Phthalate Case Study

Document cover:

The document cover was designed by Katherine Loizos of SRA International, Creative Services, Cincinnati, OH. Permission was obtained to use the following images:

Human chromosome image: Elsevier Press. The source is the article published in Genomics 9(4) by Ward, DC and Baldini, A in the article, *in situ* hybridization banding of human chromosomes with Alu-PCR products: a simultaneous karyotype for gene mapping studies, pp. 770-774. Copyright Elsevier (1991).

Microarray experiment output image: Poirazi Laboratory, Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology - Hellas, Crete, Greece.

Gene network image: Justen Andrews, Indiana University. The data for the gene network image are from the paper: Costello JC, Dalkilic MM, Beason SM, et al. (2009). Gene networks in Drosophila melanogaster: integrating experimental data to predict gene function. Genome Biol 2009 Sep 16;10(9):R97. [Epub ahead of print]

EPA/600/R-09/028F
September 2009

An Approach to Using Toxicogenomic Data In U.S. EPA Human Health Risk Assessments: A Dibutyl Phthalate Case Study

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Preferred citation:

U.S. Environmental Protection Agency (EPA). (2009) An approach to using toxicogenomic data in U.S. EPA human health risk assessments: a dibutyl phthalate case study. National Center for Environmental Assessment, Washington, DC; EPA/600/R-09/028F. Available from the National Technical Information Service, Springfield, VA, and online at http://www.epa.gov/ncea.

CONTENTS

LIST OF TABLES vii
LIST OF FIGURES ix
LIST OF ABBREVIATIONS AND ACRONYMS xi
PREFACE xiv
AUTHORS, CONTRIBUTORS, AND REVIEWERS xv
ACKNOWLEDGMENTS xvi

1. EXECUTIVE SUMMARY 1-1
   1.1. APPROACH 1-1
   1.2. DBP CASE STUDY 1-2
   1.3. RECOMMENDATIONS 1-5
   1.4. RESEARCH NEEDS 1-7

2. INTRODUCTION 2-1
   2.1. PURPOSE 2-1
   2.2. REPORT OVERVIEW 2-3
   2.3. USE OF TOXICOGENOMICS IN RISK ASSESSMENT 2-4
      2.3.1. Definitions 2-4
      2.3.2. Current Efforts to Utilize Toxicogenomic Data in Risk Assessment 2-7
         2.3.2.1. Toxicogenomics Informs TD 2-7
         2.3.2.2. Toxicogenomics Informs Dose-Response 2-9
         2.3.2.3. Toxicogenomics Informs Interspecies Extrapolations 2-10
         2.3.2.4. Toxicogenomics Informs Intraspecies Variability 2-11
         2.3.2.5. TK/TD Linkages Informed by Toxicogenomic Data 2-11
         2.3.2.6. Toxicogenomic Activities at the U.S. Food and Drug Administration (FDA) 2-12
         2.3.2.7. Toxicogenomic Activities at EPA 2-14
         2.3.2.8. Toxicogenomic Activities at Other Agencies and Institutions 2-16
      2.3.3. Current Challenges and Limitations of Toxicogenomic Technologies 2-18
   2.4. INTRODUCTION TO THE CASE STUDY 2-19
      2.4.1. Project Team 2-19
      2.4.2. Chemical Selection 2-19
         2.4.2.1. Six Candidate Chemicals 2-20
         2.4.2.2. Selection of the Case-Study Chemical 2-21
      2.4.3. Case-Study Scope 2-23

3. DBP CASE-STUDY APPROACH AND EXERCISE 3-1
   3.1. EVALUATING THE EXTERNAL REVIEW DRAFT OF THE IRIS TOXICOLOGICAL REVIEW (TOX REVIEW) OF DBP 3-1
   3.2. CONSIDERATION OF RISK ASSESSMENT ASPECTS THAT TOXICOGENOMIC DATA MAY ADDRESS 3-3
      3.2.1. Informing TK 3-7

CONTENTS (continued)

3.2.1.1. Identification of Potential Metabolic and Clearance Pathways 3-7
3.2.1.2. Selection of Appropriate Dose Metrics 3-8
3.2.1.3. Intra- and Interspecies Differences in Metabolism 3-8
3.2.1.4. TK/TD Linkages and Feedback 3-9
3.2.1.5. Research Needs for Toxicogenomic Studies to Inform TK 3-9
3.2.1.6. DBP Case Study: Do the Available Toxicogenomic Data Inform TK? 3-10
3.2.2. Informing Dose-Response 3-13
3.2.2.1. DBP Case Study: Do the Toxicogenomic Data Inform Dose-Response? 3-13
3.2.3. Informing TD 3-14
3.2.3.1. General Considerations: TD Portion of Mechanisms of Action and MOAs 3-14
3.2.3.2. DBP Case Study: MOAs for Male Reproductive Developmental Effects 3-14
3.3. IDENTIFYING AND SELECTING QUESTIONS TO FOCUS THE DBP CASE STUDY 3-17

4. EVALUATION OF THE REPRODUCTIVE DEVELOPMENTAL TOXICITY DATA SET FOR DBP 4-1
4.1. CRITERIA AND RATIONALE FOR INCLUSION OF TOXICOLOGY STUDIES IN THE EVALUATION 4-2
4.2. REVIEW OF THE TOXICOLOGY DATA SET 4-10
4.3. UNEXPLAINED MOAs FOR DBP MALE REPRODUCTIVE TOXICITY OUTCOMES 4-21
4.4. CONCLUSIONS ABOUT THE TOXICITY DATA SET EVALUATION: DECISIONS AND RATIONALE 4-26

5. EVALUATION OF THE DBP TOXICOGENOMIC DATA SET 5-1
5.1. METHODS FOR ANALYSIS OF GENE EXPRESSION: DESCRIPTION OF MICROARRAY TECHNIQUES AND SEMI-QUANTITATIVE RT-PCR 5-1
5.1.1. Microarray Technology 5-1
5.1.2. Reverse Transcription-Polymerase Chain Reaction (RT-PCR) 5-2
5.2. REVIEW OF THE PUBLISHED DBP TOXICOGENOMIC STUDIES 5-3
5.2.1. Overview of the Toxicogenomic Studies 5-3
5.2.2. Microarray Studies 5-3
5.2.2.1. Shultz et al. (2001) 5-3
5.2.2.2. Bowman et al. (2005) 5-7
5.2.2.3. Liu et al. (2005) 5-8
5.2.2.4. Thompson et al. (2005) 5-10
5.2.2.5. Plummer et al. (2007) 5-12

CONTENTS (continued)

5.2.3. RT-PCR Studies 5-14
5.2.3.1. Barlow et al. (2003) 5-14
5.2.3.2. Lehmann et al. (2004) 5-15
5.2.3.3. Thompson et al. (2004) 5-16
5.2.3.4. Wilson et al. (2004) 5-17
5.2.4. Study Comparisons 5-18
5.2.4.1. Microarray Study Methods Comparison 5-18
5.2.4.2. RT-PCR Study Methods Comparison 5-19
5.3. CONSISTENCY OF FINDINGS 5-22
5.3.1. Microarray Study Findings 5-22
5.3.2. RT-PCR Gene Expression Findings 5-25
5.3.3. Protein Study Findings 5-25
5.3.4. DBP Toxicogenomic Data Set Evaluation: Consistency of Findings Summary 5-26
5.4. DATA GAPS AND RESEARCH NEEDS 5-30
5.5. PATHWAY ANALYSIS OF DBP MICROARRAY DATA 5-31
5.5.1. Objective of the Reanalysis of the Liu et al. (2005) Study 5-31
5.5.2. Pathway Analysis of Liu et al. (2005) Utilizing Two Different Methods to Generate Hypotheses for MOAs Underlying the Unexplained Testes Endpoints 5-32
5.5.2.1. Two Methods for Identifying Differentially Expressed Genes (DEGs) 5-34
5.5.2.2. Pathway Analysis 5-37
5.6. CONCLUSIONS 5-46

6. EXPLORATORY METHODS DEVELOPMENT FOR ANALYSIS OF GENOMIC DATA FOR APPLICATION TO RISK ASSESSMENT 6-1
6.1. OBJECTIVES AND INTRODUCTION 6-1
6.2. PATHWAY ANALYSIS AND GENE INTERACTIONS AFTER *IN UTERO* DBP EXPOSURE 6-2
6.2.1. Pathway Activity Approach 6-2
6.2.1.1. Significance Analysis of Pathway Activity Levels 6-3
6.2.1.2. Pathway Activity Analysis 6-4
6.2.2. Developing a Temporal Gene Network Model 6-9
6.3. EXPLORATORY METHODS: MEASURES OF INTERSPECIES (RAT-TO-HUMAN) DIFFERENCES IN TOXICODYNAMICS 6-12
6.4. CONCLUSIONS 6-19

7. CONCLUSIONS 7-1
7.1. APPROACH FOR EVALUATING TOXICOGENOMIC DATA IN CHEMICAL ASSESSMENTS 7-1
7.2. DBP CASE-STUDY FINDINGS 7-4

CONTENTS (continued)

7.2.1. MOA Case Study Question: Do the DBP Genomic Data Inform Mechanism(s) of Action and MOA(s)? 7-5
7.2.2. Interspecies MOA Case Study Question: Do the DBP Genomic Data Inform Interspecies Differences in TD? 7-8
7.2.3. Application of Genomic Data to Risk Assessment: Exploratory Methods and Preliminary Results 7-9
7.2.4. Application of Genomic Data to Risk Assessment: Using Data Quantitatively 7-10
7.3. LESSONS LEARNED 7-12
7.3.1. Research Needs 7-13
7.3.1.1. Data Gaps and Research Needs: DBP 7-13
7.3.1.2. Research Needs for Toxicity and Toxicogenomic Studies for Use in Risk Assessment 7-14
7.3.2. Recommendations 7-16
7.3.3. Application of Genomic Data to Risk Assessment: Future Considerations 7-18

APPENDIX A: SUPPORTING TABLES FOR CHAPTER 5 ..... A-1

APPENDIX B: SUPPORTING TABLES AND FIGURES FOR CHAPTER 6 ..... B-1

APPENDIX C: QUALITY CONTROL AND ASSURANCE C-1

GLOSSARY G-1

REFERENCES R-1

LIST OF TABLES

2-1. Information available July 2005 on the selection criteria for the six candidate chemicals affecting the androgen-mediated male reproductive developmental toxicity pathway 2-22

4-1. Studies with exposures during development that have male reproductive outcomes (limited to reproductive organs and/or reproductive function) and were considered adequate for reference value determination 4-4

4-2. Reporting and study size characteristics of male reproductive studies following *in utero* exposure to DBP 4-12

4-3. Life stage at observation for various male reproductive system outcomes assessed in studies of developmental exposure to DBP 4-15

4-4. Age of assessment for individual endpoints across studies of the male reproductive system following developmental exposure to DBP 4-16

4-5. Incidence of gross pathology in F1 male reproductive organs in one continuous breeding study with DBP 4-22

4-6. Evidence for MOAs for the observed effects in the male reproductive system after *in utero* DBP exposure 4-24

5-1. Study comparisons for the toxicogenomic data set from male tissues after *in utero* DBP exposure 5-4

5-2. Lehmann et al. (2004) DBP dose-response gene expression data measured by RT-PCR showing statistically significant changes from control 5-16

5-3. Method comparisons for DBP microarray studies 5-19

5-4. Method comparisons among the RT-PCR DBP studies 5-20

5-5. Evaluation of the published protein studies after DBP *in utero* exposure (testes only) 5-27

5-6. Common pathways between the REM and SNR analyses of differentially expressed genes (DEGs) after *in utero* DBP exposure from the Liu et al. (2005) data 5-39

5-7. Genes involved in cholesterol biosynthesis/metabolism that were identified by both the REM and SNR analyses of Liu et al. (2005) 5-42

6-1. The KEGG pathways ordered based on their *p*-value for pathway activity 6-7

LIST OF TABLES (continued)

6-2. The amino acid sequence similarity of the enzymes in the steroidogenesis pathway between rat and human 6-17

7-1. DBP dose-response progression of statistically significant events illustrated with a subset of precursor event data (steroidogenesis gene expression, T expression) and *in vivo* endpoints with the reduced T MOA 7-11

7-2. Research needs for toxicogenomic studies to be used in risk assessment 7-15

7-3. Research needs for toxicity studies for utilizing toxicogenomic and toxicity data together in risk assessment 7-17

LIST OF FIGURES

2-1. The relationship between the project process, goals, and products for the development of an approach and case study for the use of toxicogenomic data in risk assessment 2-2

2-2. Androgen-mediated male reproductive developmental toxicity pathway 2-20

3-1. DBP case-study approach for evaluating toxicogenomic data for a health assessment 3-2

3-2. Exposure response array for candidate endpoints and PODs for RfD derivation presented in the external review draft IRIS Tox Review for DBP (U.S. EPA, 2006a) 3-4

3-3. Potential uses of toxicogenomic data in chemical screening and risk assessment 3-5

3-4. Potential uses of toxicogenomic data in understanding mechanisms of action 3-6

3-5. The fetal Leydig cell in the fetal testis 3-12

3-6. Approach to utilizing toxicity and toxicogenomic data for identifying affected pathways and candidate modes and mechanisms of action 3-15

3-7. The proposed DBP mechanism of action for the male reproductive developmental effects 3-16

4-1. The process for evaluating the male reproductive developmental toxicity data set for low-dose and low-incidence findings 4-3

4-2. The process for evaluating the MOA for individual male reproductive system outcomes following developmental DBP exposure 4-23

5-1. Venn diagram illustrating similarities and differences in significant gene expression changes observed in three recent microarray studies of the testes: Thompson et al. (2005), Plummer et al. (2007), and Liu et al. (2005) 5-24

5-2. Summary of DBP-induced changes in fetal gene and protein expression 5-29

5-3. Schematic of the two analysis methods (REM and SNR) for identifying differentially expressed genes and subsequent pathway analysis using GeneGo 5-33

5-4. Heat map of 1,577 DEGs from SNR analysis method 5-37

5-5. Mapping the Liu et al. (2005) data set onto the canonical androstenedione and testosterone (T) biosynthesis and metabolism pathway in MetaCore (GeneGo) 5-43

LIST OF FIGURES (continued)

6-1. An illustration of the adapted version of pathway activity level analysis for the tryptophan metabolism pathway, a nonactive pathway for DBP 6-5

6-2. Metabolic pathway network for DBP (Liu et al., 2005 data) using the pathway activity method and the KEGG database 6-8

6-3. The relationship between differential expression of individual genes and pathway activity using the Liu et al. (2005) DBP data 6-10

6-4. A gene network for DBP data of Liu et al. (2005) generated using Ingenuity Pathway Analysis (IPA) 6-11

6-5. A temporal gene network model created by IPA from the informative gene list based on time-course data after *in utero* DBP exposure 6-13

6-6. The phylogenetic relations among eight organisms based on enzyme presence, for the biosynthesis of steroids pathway, and based on information available on the NCBI taxonomy website (Sayers et al., 2008) 6-15

7-1. Approach for evaluating and incorporating genomic data into future chemical assessments 7-2

LIST OF ABBREVIATIONS AND ACRONYMS

Please note that most gene and protein name abbreviations are not included in this list because of the large number of genes and proteins described in the report. The gene and protein names have been standardized using information from the Rat Genome Project.

ADH	alcohol dehydrogenase
ADME	absorption, distribution, metabolism, and excretion
AGD	anogenital distance
AMH	anti-mullerian hormone
ANOVA	analysis of variance
AR	androgen receptor
BBDR	biologically based dose-response
BBP	butyl benzyl phthalate
BMD	benchmark dose
BMDL	benchmark dose lower confidence limit
BPA	bisphenol A
cDNA	complementary DNA
CNPs	copy-number polymorphisms
DBP	dibutyl phthalate
DEG	differentially expressed gene
DEHP	di-(2-ethylhexyl) phthalate
DEP	diethyl phthalate
DMP	dimethyl phthalate
DOTP	diocytyl tere-phthalate
DPP	dipentyl phthalate
EDC	endocrine disrupting chemical
EPA	Environmental Protection Agency
ER	estrogen receptor
ESTs	expressed sequence tags
FDA	Food and Drug Administration
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GD	gestation day
GO	Gene Ontology
GSH	glutathione

LIST OF ABBREVIATIONS AND ACRONYMS (continued)

HESI	Health and Environmental Sciences Institute
ILSI	International Life Sciences Institute
IPA	Ingenuity® Pathway Analysis
IPCS	International Programme on Chemical Safety
IRIS	Integrated Risk Information System
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Leydig cell
LMW	low molecular weight
LOAEL	lowest-observed-adverse-effect level
LOEL	lowest-observed-effect level
MAPK/ERK	mitogen-activated protein kinase/extracellular signal-regulated kinase
MAQC	MicroArray Quality Control
MAS	microarray suite
MBP	monobutylphthalate
MMP	matrix metalloproteinase
MOA	mode of action
mRNA	messenger RNA
NCCT	National Center for Computational Toxicology
NCEA	National Center for Environmental Assessment
NIEHS	National Institute of Environmental Health Sciences
NOAEL	no-observed-adverse-effect level
NOEL	no-observed-effect level
NRC	National Research Council
NTP	National Toxicology Program
PA	pathway activity
PBPK	physiologically-based pharmacokinetic
PCA	principal component analysis
PCR	polymerase chain reaction
PFOA	perfluorooctanoic acid
PND	postnatal day
POD	point of departure
PPAR	peroxisome proliferator-activated receptor
PPS	preputial separation

LIST OF ABBREVIATIONS AND ACRONYMS (continued)

RA	risk assessment
RACB	reproductive assessment by continuous breeding
RfD	reference dose
RT-PCR	reverse transcription-polymerase chain reaction
SD	Sprague-Dawley
SLR	signal log ratio
SNPs	single nucleotide polymorphisms
SNR	signal-to-noise ratio
SPC	Science Policy Council
STAR	Science to Achieve Results
T	testosterone
TD	toxicodynamics
TF	transcription factor
TK	toxicokinetics
Tox Review	Toxicological Review
UF_H	intraspecies uncertainty factor
UMDNJ	University of Medicine and Dentistry of New Jersey
VLI	valine, leucine, isoleucine
WD	Wolffian duct
WOE	weight-of-evidence

PREFACE

The U.S. Environmental Protection Agency (EPA) is interested in developing methods to use genomic data most effectively in risk assessments performed at EPA. The National Center for Environmental Assessment (NCEA) within the Office of Research and Development (ORD) prepared this document for the purpose of describing and illustrating an approach for using toxicogenomic data in risk assessment. The approach and dibutyl phthalate (DBP) case study described in this document were developed by a team of scientists at EPA laboratories and centers, and outside organizations including The Hamner Institutes for Health Sciences, the National Institute of Environmental Health Sciences (NIEHS), and the EPA National Center for Environmental Research (NCER) Science to Achieve Results (STAR) Environmental Bioinformatics and Computational Toxicology (Comp Tox) Center at the University of Medicine and Dentistry of New Jersey (UMDNJ) and Rutgers University. The intended audience for this document includes risk assessors as well as scientists with expertise in genomics, bioinformatics, toxicology, and statistics. The approach outlined in this document is expected to be useful to EPA risk assessors in the Integrated Risk Information System (IRIS) Program and other program offices and regions, as well as the scientific community at large. The review of the literature on the use of genomic data in risk assessment, as well as discussions of issues, recommendations, and methods for evaluating and analyzing toxicogenomic data, could be useful to scientists and risk assessors within and outside of EPA. The research needs identified in this document will be useful to scientists performing toxicology and toxicogenomic research studies for application to risk assessment. The DBP case study presented in this document is a separate activity from the IRIS DBP health assessment. The review of the literature included in this document was last updated in July 2007.

AUTHORS, CONTRIBUTORS, AND REVIEWERS

AUTHORS

Susan Euling, NCEA-W, U.S. EPA, Washington, DC (Lead)

Ioannis Androulakis, NCER STAR Bioinformatics Center, ebCTC, Rutgers/UMDNJ, Piscataway, NJ

Bob Benson, Region 8, U.S. EPA, Denver, CO

Weihsueh Chiu, NCEA-W, U.S. EPA, Washington, DC

Paul Foster, NIEHS, Research Triangle Park (RTP), NC

Kevin Gaido, U.S. FDA, Rockville, MD; formerly of *The Hamner Institutes for Health Sciences, RTP, NC

L. Earl Gray Jr., NHEERL, U.S. EPA, RTP, NC

Susan Hester, NHEERL, U.S. EPA, RTP, NC

Channa Keshava, IRIS, U.S. EPA, Washington, DC

Nagalakshmi Keshava, NCEA-W, U.S. EPA, Washington, DC

Andrea Kim, Allergan, Inc., Irvine, CA; formerly of *NCEA-W, U.S. EPA, Washington, DC

Susan Makris, NCEA-W, U.S. EPA, Washington, DC

Meric Ovacik, NCER STAR Bioinformatics Center, ebCTC, Rutgers/UMDNJ, Piscataway, NJ

Banalata Sen, NIEHS; formerly of *NCEA-RTP, U.S. EPA, RTP, NC

Chad Thompson, ToxStrategies, Katy, TX; formerly of *NCEA-W, U.S. EPA, Washington, DC

Lori White, *NIEHS; formerly of *NCEA-RTP, U.S. EPA, RTP, NC

Vickie Wilson, NHEERL, U.S. EPA, RTP, NC

CONTRIBUTORS

Stan Barone, NCEA-W, U.S. EPA, Washington, DC

Marianthi Ierapetritou, NCER STAR Bioinformatics Center, ebCTC, Rutgers/UMDNJ, Piscataway, NJ

---

*Affiliation at the time of work on this project for scientists with more than one affiliation listed.

AUTHORS, CONTRIBUTORS, AND REVIEWERS (CONTINUED)

REVIEWERS

Internal

Maureen R. Gwinn, NCEA-W, ORD, Washington, DC

Michael Hemmer, NHEERL, ORD, Gulf Breeze, FL

Nancy McCarroll, Office of Pesticide Programs (OPP), Washington, DC

Gregory Miller, Office of Children's Health Protection and Environmental Education (OCHPEE), Washington, DC

Marian Olsen, Region 2, New York, NY

Santhini Ramasamy, Office of Water (OW), Washington, DC

Jennifer Seed, Office of Pollution Prevention and Toxics (OPPT), Washington, DC

Imran Shah, National Center for Computational Toxicology (NCCT), ORD, RTP, NC

Jamie Strong, IRIS, ORD, Washington, DC

Dan Villeneuve, NHEERL, ORD, Duluth, MN

External

Jeanne Manson (chair), Exponent, Inc.

Robert Chapin, Pfizer, Inc.

Julia Gohlke, NIEHS

Hisham Hamadeh, Amgen, Inc.

Poorni Iyer, California EPA (Note: Dr. Iyer conducted the peer review as a private consultant and not as a representative of California EPA)

ACKNOWLEDGMENTS

This project was funded by NCEA and the EPA NCCT's Research Program under their new starts grants. We thank the outside partners, NIEHS and The Hamner Institutes for Health Sciences, for allowing team members at these institutions to work on this project. Some of the work described was performed at the STAR Bioinformatics Center at UMDNJ and Rutgers University that is supported by the grant R832721 from the EPA's STAR program. We gratefully acknowledge Dr. Kevin Gaido for providing the data from the Liu et al. (2005) study performed in his laboratory at The Hamner Institutes for Health Sciences, Terri Konoza of NCEA for her detailed editorial contribution to this document, and Sarah Burgess-Herbert for her thoughtful review of Chapter 6.