DRAFT
DO NOT CITE OR QUOTE

EPA/600/R-09/028A
May 2009
External Review Draft

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

NOTICE

THIS DOCUMENT IS AN EXTERNAL REVIEW DRAFT. It has not been formally released by the U.S. Environmental Protection Agency and should not at this stage be construed to represent U.S. EPA policy. It is being circulated for comment on its technical policy implications.

Please note gene and protein names in this document have been standardized using information from the Rat Genome Project.

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

DISCLAIMER

This document is a draft for review purposes only and does not constitute U.S. EPA policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS AND ACRONYMS

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
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
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 for Environmental Health Sciences
NOAEL	no-observed-adverse-effect level
NOEL	no-observed-effect level
NRC	National Research Council
NTP	National Toxicology
OPA	overall 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
RACB	reproductive assessment by continuous breeding
RfD	reference dose
RT-PCR	reverse-transcriptase polymerase chain reaction
SD	Sprague-Dawley
SLR	signal log ratio

LIST OF ABBREVIATIONS AND ACRONYMS (continued)

SNP	single nucleotide polymorphism
SNR	Signal-to-Noise Ratio
SPC	Science Policy Council
STAR	Science to Achieve Results
T	testosterone
TD	toxicodynamic
TF	transcription factor
TgX in RA	Toxicogenomics in Risk Assessment
TK	toxicokinetic
UF_H	intraspecies uncertainty factor
UMDNJ	University of Medicine and Dentistry of New Jersey
U.S. EPA	United States Environmental Protection Agency
VLI	valine, leucine, isoleucine
WD	Wolffian duct
WOE	weight-of-the-evidence

PREFACE

The United States Environmental Protection Agency (U.S. EPA) is interested in developing methods to use genomic data most effectively in risk assessments performed at the U.S. EPA. The National Center for Environmental Assessment (NCEA) 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 U.S. EPA laboratories and centers, and outside organizations including The Hamner Institute (formerly CIIT), the National Institute for Environmental Health Sciences (NIEHS), and the U.S. EPA Science to Achieve Results (STAR) Bioinformatics 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 U.S. 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 U.S. 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 2008.

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, National Institutes of Environmental Health Sciences (NIEHS), Research Triangle Park (RTP), NC

Kevin Gaido, 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, Allergen; formerly of NCEA-W

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, Research Triangle Park, NC

Chad Thompson, 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

REVIEWERS

Internal

Maureen R. Gwinn, NCEA-W, ORD

Michael Hemmer, NHEERL Gulf Breeze, ORD

Nancy McCarroll, OPP

Gregory Miller, OCHPEE

Marian Olsen, R2

Santhini Ramasamy, OW

Jennifer Seed, OPPTS

Imran Shah, NCCT, ORD

Jamie Strong, IRIS, ORD

Dan Villeneuve, NHEERL Duluth, ORD

ACKNOWLEDGMENTS

This project was funded by U.S. EPA's National Center for Environmental Assessment (NCEA) and U.S. EPA's National Center for Computational Toxicology's (NCCT) Research Program under their new starts grants. We thank the outside partners, NIEHS and The Hamner Institute, 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 U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program. We gratefully acknowledge Dr. Kevin Gaido for providing raw data from the Liu et al. (2005) study.

1. EXECUTIVE SUMMARY

We developed a systematic approach for evaluating and utilizing toxicogenomic data in health assessment. This document describes this approach and describes a case study we conducted on dibutyl phthalate (DBP) to illustrate and refine the proposed approach. DBP was selected for the case study because it has a relatively large genomic data set and phenotypic anchoring of certain gene expression data to some male reproductive developmental outcomes. A U.S. Environmental Protection Agency (U.S. EPA) Integrated Risk Information System (IRIS) assessment of DBP is ongoing but the case study described here is a separate endeavor, with distinct goals.

Toxicogenomics is the application of genomic technologies (e.g., transcriptomics, genome sequence analysis) to study effects of environmental chemicals on human health and the environment. Currently, the U.S. EPA provides no guidance for incorporating genomic data into risk assessments of environmental agents. However, the U.S. EPA's Science Policy Council (SPC) has developed guidance regarding other aspects of using microarray data, entitled *Interim Guidance for Microarray-Based Assays: Data Submission, Quality, Analysis, Management, and Training Considerations*. In this document, we review some of the recent and ongoing activities regarding the use of genomic data in risk assessment, inside and outside of the U.S. EPA.

1.1. APPROACH

Genomic data have the potential to inform mechanism of action, inter- and intra-species toxicodynamic differences, exposure assessment, toxicokinetics, and dose-response assessment. Our strategy for evaluating genomic data for risk assessment was to design a systematic approach to evaluating the genomic data set for a particular chemical that is flexible enough to accommodate different health and risk assessment practices. The first step of the approach is to evaluate the available genomic data set for their application to a broad range of information types (e.g., mode of action [MOA], toxicokinetics [TK], interspecies variability) useful to risk assessment as well as the steps of health assessment (e.g., hazard characterization, dose-response assessment). Through this iterative process, the potential use of the available genomic data is determined. As part of this scoping step, a review of all available data sets (e.g., epidemiology, toxicology, genomics) further determines the potential applications of the genomic data. The

toxicity, human, and toxicogenomic data sets are considered together to determine the relationship or phenotypic anchoring between gene and pathway changes and health or toxicity outcomes. As a result of the scoping step, questions are posed to direct the genomic data set evaluation.

The next steps include detailed evaluations, directed by the formulated questions, of the outcome (either toxicity or human health outcomes of interest) and the toxicogenomic data set. For example, when data are available to inform mechanism or mode of action, the toxicogenomic and toxicity data sets can be evaluated together, relating the affected endpoints (identified in the toxicity data set evaluation) to the pathways (identified in the toxicogenomic data set evaluation) to establish or formulate hypotheses about the mechanism or MOA. In addition to informing the mechanism of action and the MOA, genomic data also have the potential to inform inter- and intraspecies toxicodynamic differences, toxicokinetics, and dose-response assessment depending on the genomic study design (e.g., species, organ, single dose vs. multiple doses, genomic method) of the available data. The approach also includes new analyses of the genomic data for the purpose of risk assessment when data are available and different analyses could address questions relevant to the risk assessment.

1.2 DBP CASE STUDY

For the DBP case study example, consideration of risk assessment information and steps was accomplished in two parallel processes. We took advantage of the DBP IRIS assessment external review draft, which summarized data sets and identified data gaps. We asked whether the genomic data set could inform any of these data gaps. In parallel, the DBP genomic data set was considered in light of all risk assessment aspects that these data might inform. As a result of following these two processes, we posed two specific case-study questions:

1) *Do the toxicogenomic data inform the mechanism and/or MOA for DBP?; and*

2) *Do the toxicogenomic and other data better inform interspecies toxicodynamic differences?*

Additional questions were excluded because appropriate data for addressing the questions was lacking. For example, one question of great interest is *Do the toxicogenomic data inform*

dose-response? However, this question could not be addressed in this case study because there were no dose-response genomic data for DBP. Few chemicals have available dose-response genomic data; DBP is not unusual in this respect. The one DBP dose-response gene expression study, although not global, is discussed in the document. As a result of the DBP genomic data set limitations, the case study focuses on the qualitative application of genomic data to risk assessment. In addition, the exposure assessment step was not considered in this approach because the case study was performed using an IRIS chemical assessment model.

For Case Study Question 1, we found that the DBP toxicogenomic data do inform the mechanism of action and possibly the MOA. There is good evidence in the published literature that a number of the gene expression changes observed in genomic studies are "phenotypically anchored" (i.e., in the causal pathway) for a number of the male reproductive developmental outcomes observed after in utero DBP exposure in the rat. The available genomic and other gene expression data, hormone measurement data, and toxicity data for DBP are instrumental in the establishment of two of its MOAs: (1) a decrease in fetal testicular testosterone (T), and (2) a decrease in Insulin-like 3 (*Insl3*) expression. A decrease in fetal testicular testosterone is the MOA for a number of the male reproductive developmental effects in the rat. The genomic and single gene expression data after in utero DBP exposure identified changes in genes involved in steroidogenesis and cholesterol transport, consistent with the observed decrease in fetal testicular testosterone. Along with the decreased fetal testosterone, a decrease in *Insl3* expression is a second MOA responsible for undescended testis descent, and this MOA is well established by reverse-transcription polymerase chain reaction (RT-PCR) and in vivo toxicology data.

Evaluating genomic and toxicity data together also provides information on putative novel MOAs. A number of the DBP toxicity and toxicogenomic studies were performed in the same strain of rat, and exposed to similar doses and at similar exposure intervals, allowing for comparions across studies. In this case study, rodent reproductive developmental toxicity studies were evaluated for low incidence and low-dose findings and for the male reproductive developmental effects that currently do not have a well-established MOA. In the case study we focused on the testes outcomes because all but one of the DBP toxicogenomic studies were performed on the testes. We identified five testes endpoints without a known MOA that were pursued further in the evaluation of the toxicogenomic data set. The nine published RT-PCR and microarray studies in the rat were evaluated as part of the toxicogenomic and associated gene

expression data set to identify genes and pathways affected after in utero DBP exposure. All of the gene expression data were evaluated for consistency of findings. At the gene level, the findings from the DBP genomic studies (i.e., microarray, RT-PCR, and protein expression) were relatively highly correlated with one another in both the identification of differentially expressed genes (DEGs) and their direction of effect. The evaluation of the published toxicity and toxicogenomic studies corroborates the two known MOAs for DBP.

New pathway identification analyses were performed for one of the published microarray studies of DBP because the published studies focused primarily on pathways related to the reduced fetal testicular testosterone MOA, such as the steroidogenesis pathway. We performed new analyses of the data from a rat testes microarray study in order to identify all possible pathways significantly affected by in utero DBP exposure. Using a variety of analytical methods, pathways associated with the two known MOAs (decreased *Insl3* and fetal testicular testosterone), as well as new processes (e.g., growth and differentiation, transcription, cell adhesion) and pathways (e.g., *Wnt* signaling and cytoskeleton remodeling) not associated with either *Insl3* or steroidogenesis pathways, were identified. The newly identified putative pathways may play a role in the regulation of steroidogenesis (i.e., related to a known MOA for DBP) or, alternatively, may inform another MOA for one or more unexplained outcomes in the testes. This approach allowed us to develop hypotheses about possible DBP MOAs for some male reproductive developmental outcomes.

For Case Study Question 2, genomic data were evaluated to inform interspecies differences in the steroidogenesis pathway, relevant to the decreased fetal testicular testosterone MOA. We explored the development of new methods to evaluate interspecies TD differences. The steroidogenesis gene and pathway information for rats and humans was compared via three approaches, protein sequence similarity, pathway network similarities, and promoter region conservation, to evaluate cross-species similarity metrics. Preliminary results from all three methods suggest that steroidogenesis genes are relatively highly conserved between rats and humans. For the DBP case, we do not recommend utilizing these data to inform interspecies uncertainty because it is difficult to make unequivocal conclusions regarding a “high” vs. “low” degree of conservation for the genes in this pathway based on these data alone. With further refinement and improved data sources, these methods could potentially be applied to other chemical assessments.

New methods for evaluating microarray data for the purposes of risk assessment were explored and developed during the DBP case study. These methods include a new pathway analysis methods designed for risk assessment application that determine pathway level changes as opposed to mapping affected genes to pathways, and utilizing this method for evaluating time course microarray data. In the DBP case study, we explored the use of methods to develop a genetic regulatory network model. Preliminary results based on data from one time course study identified a temporal sequence of gene expression and pathway interactions that occur over an 18-hour interval within the critical window of exposure for DBP and testicular development effects.

1.3 RECOMMENDATIONS

In addition to following the principles of the approach (i.e., systematically consider all types of information with respect to the steps of risk assessment and evaluate genomic data and toxicity data together), several specific methodological recommendations arose from the DBP case study. Two of these recommendations are straightforward and could reasonably be performed by a risk assessor with basic genomics training:

1) *Evaluate the genomic and other gene expression data for consistency of findings across studies to provide a weight-of-the-evidence (WOE) evaluation of the affected gene expression and pathways.* Some simple methods, such as using Venn diagrams and gene-expression compilation approaches can be applied to risk assessment. When evaluating the consistency of toxicogenomic data findings, it was advantageous to include all of the available gene expression data (single gene, global gene expression, protein, RNA) because the single gene expression techniques have been traditionally used to confirm the results of global gene expression studies.

2) *Perform benchmark dose (BMD) modeling on high-quality RT-PCR dose-response studies for genes known to be in the causal pathway of a MOA or outcome of interest.* Obtaining a BMD and BMDL (benchmark dose lower confidence limit) is a useful starting point for both linear low-dose extrapolation and reference value approaches. We are not indicating which approach is appropriate to take for making predictions about the potential risk below the BMD or BMDL. "High quality" is defined in this context as a well conducted study that assessed enough animals and litters for sufficient statistical power for characterizing the mean responses and the variability (interlitter and intralitter variability).

Two additional recommendations require expertise in genomic data analysis methods to implement:

3) *Perform new analysis of toxicogenomic raw data in order to identify all affected pathways or for other risk assessment applications.* Most often, microarray studies are conducted for different purposes (e.g., basic science, pharmaceutical development). In these cases, new pathway analysis of microarray data can be potentially useful.

4) *Develop a genetic regulatory network model for the chemical of interest to define the system of interacting regulatory DNA sequences, expression of genes, and pathways for one or more outcomes of interest.* Genetic regulatory network model methods, developed as part of this case study, could be used in a risk assessment. If time-course genomic data are available, the temporal sequence of mechanistic events after chemical exposure can be defined, and the earliest affected genes and pathways, that may be define the initiating event, may be identified.

Based on these recommendations, we refined the approach that was used in the case study that can be useful for evaluating genomic data in new chemical assessments.

1.4 RESEARCH NEEDS

We identified the following research needs to improve the utility of genomic data in risk assessment:

- Perform parallel toxicity and toxicogenomic study-design characteristics (i.e., dose, timing of exposure, organ/tissue evaluated) to obtain comparable results to aid our understanding of the linkage between gene expression changes and phenotypic outcomes;
- Collect exposure time-course microarray data to develop a regulatory network model;
- Generate TK data in a relevant study (time, dose, tissue), and obtain a relevant internal dose measure to derive the best internal dose metric;
- Test multiple doses in microarray studies in parallel with phenotypic anchoring in order to relate dose, gene expression response, and in vivo response;
- Continue further development of bioinformatic methods for analyzing genomic data for use in health and risk assessments.

As a result of considering how to best use genomic data in risk assessment, we identified a number of issues for future consideration. As more and various types of genomic studies are performed, genomic data will likely inform multiple steps of the risk assessment process beyond

its use to inform MOA. To facilitate the advancement of the use of genomics in risk assessment, first, we need approaches to utilize genomic data quantitatively in risk assessment, for application to dose-response, intraspecies variability, and TK. Second, analytical methods tailored to use in risk assessment are needed. Methods development work, some initiated in this project, has made significant progress in adapting bioinformatic methods used for hypothesis generation to the express purpose of utilizing genomic data for risk assessment. However, continued effort, with input from statistical modeling and biology experts, is required to validate and test these methods, and develop newer methods. Third, training risk assessors in analysis methods of genomic data would assist the U.S. EPA in being able to both analyze complex, high-density data sets and to perform new analyses when necessary.

Finally, some of the current issues in utilizing genomic data in health and risk assessment are not unique to genomic data but apply to precursor event information in general. Two of these issues are (1) defining adversity and (2) establishing biological significance, in the case of genomics, of gene expression changes or a pattern of gene expression. The design and performance of appropriate studies, with both genomic and toxicity components, are needed to address these two important issues.

As far as we know, this is the first systematic approach for using genomic data in health assessment at U.S. EPA. We believe that this document can serve as a template that risk assessors can use when considering a large range of potential applications, issues, and methods to analyze genomic data that can be applied to future assessments. This approach advances efforts in the regulatory and scientific communities to devise strategies for using genomic data in risk assessment, and it is consistent with the pathway-based risk assessment vision for the future outlined in the National Research Council (NRC) *Toxicity Testing in the 21st Century*. We also anticipate that the research needs and future considerations described herein will advance the design of future toxicogenomic studies for application to risk assessment, and thus, benefit the bioinformatic, toxicogenomic, and risk-assessment communities.

2. INTRODUCTION

2.1. PURPOSE

Currently, the U.S. EPA provides no guidance for incorporating genomic data into risk assessments. The project addressed the question of how the available toxicogenomic data may be best used to improve U.S. Environmental Protection Agency (U.S. EPA) human health risk assessments. Specific questions motivating the project include

- *Could toxicogenomic data inform one or more steps (e.g., dose-response) in the risk assessment process?;*
- *How could current issues (e.g., reproducibility, variability in response) with the use of genomic technologies, particularly microarrays, be taken into account in the evaluation of genomic data?; and*
- *How could toxicogenomic data be used in conjunction with other types of information?*

After considering the overarching questions listed above, we chose to focus on developing an approach for using toxicogenomic data in U.S. EPA human health assessments because a practical approach would have broad application to risk assessment methods. The specific goals of this methods development project were to

- *Develop a systematic approach that allows the risk assessor to utilize the available toxicogenomic data in chemical-specific health risk assessments performed at U.S. EPA; and*
- *Perform a case study to illustrate the approach.*

2.2. REPORT OVERVIEW

This report describes an approach to evaluating toxicogenomic data for use in risk assessment and a case study for the chemical DBP. The approach principles includes examination of genomic and toxicity datasets, defining a set of questions to direct the evaluation, and performing new analyses of genomic data, when available. The DBP case study example

focuses on male reproductive developmental effects and the qualitative use of genomic data in risk assessment.

Currently, EPA provides no guidance for evaluating and incorporating genomic data into risk assessment. In the approach described in this document, the genomic data are evaluated for their application to a broad range of information types useful in risk assessment; and both the genomic and toxicity datasets are considered together to determine the relationship between genomic changes and health outcomes and inform the mechanism of toxicity. The document includes the development of exploratory methods and preliminary results from genomic data analysis. In addition, recommendations, research needs, and potential future directions arre identified.

This chapter (Chapter 2) includes a focused review of the history and current use of genomic data in risk assessment and the rationale for selecting DBP as the case-study chemical. Chapter 3 presents the approach that we developed for use of toxicogenomic data in risk assessment used for the DBP case study. This includes discussions of the various steps of the approach that can be used in future assessments. Chapters 4−6 present the DBP case study data evaluations and analyses. Chapter 4 presents the toxicology data set evaluation, Chapter 5 presents the toxicogenomic data set evaluation, and Chapter 6 presents the new analyses of some of the DBP genomic studies, and exploratory methods that were developed. Supplemental material for the work described in Chapters 5 and 6 are presented in Appendices A and B. Chapter 7 presents the case study conclusions including a refined approach for evaluating genomic data for risk assessment, research needs, and future considerations.

The audience for the various chapters varies because of the highly technical nature of some of the work performed. Risk assessors will benefit from Chapters 2-5 and Chapter 7 because it describes the approach and case study evaluations based on the published literature only. Bioinformaticians and risk assessors trained in analyzing microarray data will find the descriptions of the pathway-analysis methods and the development of new methods in Chapter 6 useful. Risk assessors and scientists performing toxicology and toxicogenomic research, inside and outside of the U.S. EPA, will benefit from the refined approach to using genomic data in U.S. EPA risk assessment and research needs presented in Chapter 7.

2.3. USE OF TOXICOGENOMICS IN RISK ASSESSMENT

Recent and ongoing activities regarding the use of genomic data in risk assessment, inside and outside of U.S. EPA, are reviewed below.

2.3.1. Definitions

Toxicogenomics is a fairly new field that studies the global expression of genes, proteins, or the concentration or relative abundance of small molecular weight metabolites after exposure to a toxic agent in order to characterize responses. Such responses are considered more sensitive and precursor in nature because the techniques measure molecular responses on a near-global scale. The techniques to generate toxicogenomic data include DNA sequencing, transcriptomics, proteomics, and metabolomics. These techniques are near-global because of annotation limitations or detection limitations.

Transcriptomics, through the use of microarrays, is a powerful tool for investigating the expression levels of thousands of genes or sometimes a complete genome, following exposure to toxicants. The use of microarrays to study gene expression profiles from tissues, organs, or cells began in 1995 (Lobenhofer et al., 2001). Microarray information is different from other types of data used in toxicology for a number of reasons, largely due to the global nature of the gene expression data. Unlike single-gene-expression data that use specific methods, such as northern blots and real-time reverse transcription-polymerase chain reaction (RT-PCR) to evaluate individual genes, microarrays provide a nearly global (i.e., not all genes are currently annotated and have expressed sequence tags [ESTs]) transcriptional profile of a cell or tissue. Thus, each experiment generates a large amount of data. Analyzing and interpreting the quantity and complex patterns of data requires expertise in bioinformatics.

The term omics (referring to terms ending with the suffix ‘omics) is a broad discipline of science and engineering for analyzing the total (“om”) or global interactions within a biological system by utilizing the various genomic, proteomic, and metabolomic techniques. These include genomics, proteomics, metabolomics, etc. The main focus is on (1) mapping information objects such as genes and proteins, (2) finding interaction relationships among the objects, and (3) engineering the networks and objects to understand and manipulate the regulatory mechanisms (For more background information about ‘omics see www.omics.org).

The scientific community has a range of definitions for the terms genomics and toxicogenomics. Toxicogenomics refers to a set of technologies for assessing the genome, transcriptome, proteome, and metabolome gene products after toxic agent exposure. In this document, we use definitions of the toxicogenomic terms that are consistent with the National Research Council (NRC) report entitled, "*Applications of Toxicogenomic Technologies to Predictive Toxicology and Risk Assessment*" (NRC, 2007a). Genomics is the study of the genome and includes genome sequencing and genotype analysis techniques (e.g., polymorphism identification). U.S. EPA's Science Policy Council (SPC) (2002) defines genomics as "the study of all the genes of a cell, or tissue, at the DNA (genome), messenger RNA (mRNA; transcriptome), or protein (proteome) levels." One goal of toxicogenomic studies is to link genomic changes with adverse phenotypic effects/outcomes determined histopathologically or clinically.

Genetic polymorphisms are included in the definition of genomic techniques. Some microarrays have been designed to detect single nucleotide polymorphisms (SNPs) and copy-number polymorphisms (CNPs; Buckley et al., 2005). Polymorphism analysis can be used qualitatively and quantitatively to assess risks to various subpopulations as well as provide insights to mechanistic pathways (Guerreiro et al., 2003; Shastry, 2006). Transcriptomics measures global mRNA expression (NRC, 2007a). The transcriptomic technology with the greatest history and success are microarrays. It is a tool used to understand specific genes and pathways involved in biological processes. Underlying the use and interpretation of these technologies is the assumption that genes exhibiting a similar expression pattern may be functionally related and under the same genetic control. Genes that are annotated as well as those that are not (i.e., ESTs) are included in microarray analysis. Global gene analysis provides information about the effect of a chemical on toxicity pathways, defined as "A series of biochemical and physiological changes that occur after chemical interaction at the target site that are linked to the adverse outcome" (U.S. EPA, 2004b). Common technologies for genome-wide or high-throughput analysis of gene expression are complementary DNA (cDNA) microarrays and oligo-microarrays, cDNA-amplified fragment length polymorphism, and serial analysis of gene expression.

Proteomics is the study of proteins in an organism (NRC, 2007a). It involves the study of the proteins: specifically, their expression, their structural status (e.g.,

phosphorylated/dephosphorylated), their functional states (i.e., activity specificity and activity level), and their interactions with other cellular components—all as a function of time and response to intrinsic and extrinsic factors (Pandey and Mann, 2000). Thus, proteomics offers the ability to study both changes in protein expression and protein modification in toxicity (Ekins et al., 2005; Anderson and Anderson, 1998), and, ultimately, changes in cellular function. Broadly, proteomics may be defined as "expression" (or "differential") proteomics and "functional" proteomics (Wu et al., 2002); the former relates to a differential expression of proteins among treatments or disease states, and the latter relates to protein interactions and changes in function due to posttranslational modifications or other protein-protein interactions.

Metabolomics is the study of low molecular weight (LMW) metabolic products (NRC, 2007a). Since metabolites are the final functional products of genes, a metabolomic profile can capture the most functional assessment of toxicity, among the omic technologies. Metabonomics is also the study of LMW protein. There is a subtle distinction between the two: metabolomics refers to the study of LMW molecules within cells, whereas metabonomics refers to a more systemic and complex change in tissues and body fluids (Ekins et al., 2005). For example, the toxicity of acetaminophen in rodents has been examined via metabonomics using nuclear magnetic resonance (NMR) spectroscopy to characterize changes in intact and solubilized liver tissue and blood plasma (Coen et al., 2003). Such approaches to examining toxicity can be used qualitatively to help define or refine the mode of action (MOA) of an environmental toxicant, potentially to serve as biomarkers for exposure, or, in some cases, quantitatively to represent a toxic response amenable to dose-response analysis. Due to the large size and complexity of information generated by omics technologies, bioinformatics methods for data analysis continue to be developed and refined.

In the DBP case study, the toxicogenomic and all other gene expression data were evaluated. We decided to include all the microarray studies detecting global gene expression, as well as single-gene and protein expression such as RT-PCR, northern blotting, transgene expression, and immunostaining in the evaluation of genomic data for risk assessment because these techniques provide (1) a validation method for microarray studies; (2) a larger data set of gene expression information, as there are typically a very small number of available microarray studies for a specific chemical; and (3) additional semiquantitative information such as RT-PCR and protein expression assays.

The mechanism of action is defined herein as the complete molecular sequence of events between the interaction of the chemical with the target site and observation of the outcome. Thus, the mechanism of action can include toxicokinetic (TK) and toxicodynamic (TD) steps. By contrast, "mode of action" is defined as a sequence of key events that the outcome is dependent upon. A "key event" is an empirically observable precursor step that is itself a necessary element of the MOA or is a biologically based marker for such an element (U.S. EPA, 2005).

2.3.2. Current Efforts to Utilize Toxicogenomic Data in Risk Assessment

Many of the advances in toxicogenomic technology are a result of their application within the pharmaceutical industry (Boverhof and Zacharewski, 2006). In drug discovery, genomic methods are used for assessing and predicting toxicity with the goal of selecting a drug with relatively high efficacy and low toxicity. Research and regulatory agencies are also interested in using omics-generated data and its implications. However, to date, their application has been somewhat limited due, at least in part, to a lack of available data and expertise required to analyze and interpret these data when available. Nevertheless, approaches and considerations to using toxicogenomic data sets in a risk assessment or other regulatory scenario continue to be explored (Boverhof and Zacharewski, 2006; Hackett and Lesko, 2003; Chan and Theilade, 2005; Cunningham et al., 2003; Frueh et al., 2004; Leighton, 2005; Oberemm et al., 2005; Pennie et al., 2004; Pettit et al., 2003; Reynolds, 2005; Robinson et al., 2003; Simmons and Portier, 2002; Waters and Fostel, 2004). An effort has been made to apply toxicogenomic data to the area of exposure assessment. For example, a few studies have used gene expression analysis successfully to determine occupational exposure levels (NRC, 2007a).

2.3.2.1. *Toxicogenomics Informs Mode of Action (MOA)*

Genomic data have been used in risk assessment to provide information about the mode's and mechanism's action. For example, toxicogenomic data can be used to complement other in vitro and in vivo toxicology data. A number of studies have used microarrays to identify patterns of gene expression following chemical exposures (Ellinger-Ziegelbauer et al., 2005; Moggs et al., 2004; Lobenhofer et al., 2001). Further, some studies have found common patterns of gene expression for specific groups of chemicals (Naciff et al., 2005; Hamadeh et al., 2002a).

Hamadeh et al. (2002) performed microarray analysis of liver tissue from animals exposed to four different chemicals: the pharmaceutical peroxisome proliferators clofibrate, Wyeth 14,643, gemfibrozil, and the CYP2B inducer phenobarbital. The three peroxisome proliferators gave similar patterns of gene expression indicating a common MOA; whereas, the gene expression pattern for phenobarbital was distinct from the three peroxisome proliferators. Naciff et al. (2005) studied the transcriptional profile in the testis following exposure to three estrogen agonists, 17α-ethynyl estradiol, genistein, or bisphenol A (BPA), which have been shown to bind to the estrogen receptor (ER) with different affinities (e.g., BPA binds most weakly). A common group of 50 genes, whose expression was changed in the same direction, was identified among the three estrogen agonists. Dose-response studies were performed, and the gene expression changes were also associated with dose (i.e., lower dose, lower gene expression) among these 50 genes for each of the three chemicals. Both of these laboratory groups found differences in gene expression patterns depending on the duration of exposure (Hamadeh et al., 2002), the organ (Naciff et al., 2005, 2002), or the life stage of exposure (Naciff et al., 2003, 2002). Recently, in addition to gene patterns and chemical signatures, Tilton et al. (2008) have identified an alternative mechanism for hepatic tumor promotion by perfluorooctanoic acid (PFOA) in rainbow trout. Using gene expression profiles, those study authors have demonstrated a novel mechanism involving estrogenic signaling for the tumor promotion activity of PFOA. In their study, tumor promotion was not related to the function of PFOA as a peroxisome or peroxisome proliferator-activated receptor alpha (PPARα) agonist, but it is phenotypically linked to estrogenic gene signatures in trout liver.

The use of omics data, particularly "gene expression signatures" or "fingerprints," to make predictions about the toxicity of a chemical based upon gene expression patterns for a given MOA class is not always straightforward. Although peroxisome proliferators may exhibit a similar gene expression signature, some chemicals (e.g., PFOA) may exert effects through multiple mechanisms. In this regard, it may be possible to be misled by the presence or absence of certain signatures, or to focus on a subset of genes in the overall signature pattern. However, the Tilton et al. (2008) study is a good example of the power of genomic signatures to identify additional MOAs.

2.3.2.2. *Toxicogenomics Informs Dose-Response*

As noted previously, most examples of the use of toxicogenomic data have focused on informing hazard characterization, TD, and MOA. However, it is also important to consider whether and how toxicogenomic data can inform dose-response analysis and TK. In regards to dose-response analysis, toxicity endpoints (e.g., hepatotoxicity) will likely have characteristic genomic profiles of associated gene expression changes that can serve as fingerprints for these toxicity mechanisms (Aardema and MacGregor, 2002). Importantly, gene changes related to a toxic response may be observable at doses lower than those required to elicit more overt toxic responses and, thus, serve as sensitive precursor effects. Alternatively, such changes may occur at doses similar to those that exert more overt effects, but at much earlier time points, and, ultimately, without the need to carry through expensive chronic bioassays. While establishing such fingerprints and validating their utility for quantitative dose-response analysis is necessary for risk and safety assessment, these gene changes could aid risk assessors in choosing the most appropriate animal model for conducting toxicity studies (Aardema and MacGregor, 2002), with the likely result of reducing uncertainties inherent in risk assessment.

Recent studies on formaldehyde lend support to the notion that gene changes may be able to serve as early indicators of longer-term in vivo outcomes (Thomas et al., 2007; Andersen et al., 2008). These studies used gene ontology (GO) categorization of microarray data after chemical exposure to chemicals that cause rodent tumor formation. The study authors observed significant changes in gene expression after chemical exposure for chemicals (e.g., formaldehyde) that lead to cell proliferation and DNA repair occur at approximately the same doses associated with long-term exposure leading to observable tumor formation in rodents. The authors conclude that relevant gene changes may serve to predict the long-term outcome of bioassays. In an editorial by Daston (2008), he suggests that gene expression changes may not occur below a threshold dose for these toxic agents. Alternatively, it is possible that longer-term exposure to low doses could lead to genomic changes in the cell that are linked to toxicity; such aspects may not be captured in the small treatment group sizes in this study or under shorter durations of exposure.

Approaches are needed to use these data quantitatively for risk assessment. Studies carried out by the Hamner Institute on formaldehyde carcinogenicity mark one of the first efforts to apply toxicogenomics data quantitatively (Thomas et al., 2007). In examining the

dose-response for formaldehyde-induced gene changes in rat nasal tissue, a benchmark dose (BMD) analysis was used to identify sets of genes in GO categories often thought to be involved in the MOA of formaldehyde (Thomas et al., 2007). GO categories for DNA damage response and repair, response to unfolded proteins, and regulation of cell proliferation all had BMD values (defined as 1.349 x standard deviation of control) ranging from 5.68 to 6.76 ppm formaldehyde. The authors noted the relatively close agreement between the BMD (5.68 ppm) for the cell proliferation GO category and a previously published BMD (4.91 ppm) for cell labeling index (Schlosser et al., 2003), as well as between the BMD (6.31 ppm) for the DNA damage response GO category and a lowest-observed-adverse-effect level (LOAEL; 6 ppm) reported for DNA-protein crosslink formation (Casanova et al., 1994). Similar conclusions were drawn from a longer-term, 3 week, study by Andersen et al. (2008). Although the justification for comparing these values (e.g., a 10% increase in cell labeling vs. 1.349 x SD for cell proliferation genes) may be debated, dose-response modeling methodologies can be developed that, upon further validation, might support the modeling of toxicogenomic data for chemicals with more limited data—either for risk assessment or general screening and prioritization purposes.

2.3.2.3. *Toxicogenomics Informs Interspecies Extrapolations*

Interspecies extrapolations are comprised of TK and TD aspects. Changes in genes, proteins, or LMW molecules that are likely involved in chemical disposition (e.g., transporters, enzymes, and cofactors) can potentially inform TK extrapolations for risk assessment. For example, changes in expression of genes or proteins related to glutathione (GSH) synthesis following exposure to an environmental toxicant suggest that further consideration of GSH (including synthesis or resynthesis) may be necessary when considering dose adjustments or building physiologically based pharmacokinetic (PBPK) models. In principle, this approach has been demonstrated for the depletion and resynthesis of GSH following exposure to trichloroethylene and 1,1-dichloroethylene, albeit without toxicogenomic data per se (El-Masri et al., 1996). In this study, modeling suggests that it is important to consider GSH resynthesis when assessing the toxicity of these chemicals. Similarly, toxicogenomic data suggesting the presence of proteins in TK may inform dosimetry modeling. Additionally, toxicogenomic and proteomic data can also inform TD aspects of interspecies extrapolation. Often chemical-specific data to account for TD differences across species are not available. Toxicogenomics

data indicating distinctions in expression profiles between species may help qualitatively and quantitatively address these issues. Again, considering changes in GSH genes, differential changes across species may have implications for TD if redox status is thought to play a role in a chemical's MOA.

2.3.2.4. *Toxicogenomics Informs Intraspecies Variability*

Perhaps the most straightforward quantitative application of toxicogenomic data in risk assessment involves genetic polymorphisms. This application is also the most amenable to current risk assessment practices—specifically in handling interindividual variation in TK. Both SNPs and chromosome CNPs in genes that are important for the disposition of environmental toxicants have the potential to inform the intraspecies uncertainty factor (UF_H) applied in risk assessments. When the impact of polymorphisms on enzyme function is known, this information can either be used to characterize the difference in dose metric for a subpopulation relative to the most common alleles, or, it can be used in probabilistic assessments using Monte Carlo analysis to incorporate population variability in enzyme function and dose metric predictions. El-Masri et al. (1999) demonstrated this approach for polymorphisms in GSH transferase-1. Ultimately, polymorphisms related to TD aspects of a chemical model of action might also be incorporated into risk assessments as more sophisticated biologically based models are developed.

2.3.2.5. *Toxicokinetic/Toxicodynamic (TK/TD) Linkages Informed by Toxicogenomic Data*

Toxicogenomic data will likely play an increasing role in the modeling of systems biology for use in risk assessment (Daston, 2007; Andersen et al., 2005). To this end, understanding the impact of xenobiotics in organisms will require greater focus and understanding of the normal biological processes and compensatory mechanisms in biological systems. Ultimately, this information will improve our understanding of the shape of dose-response curves at environmentally relevant concentrations and for low-incidence adverse effects (Andersen et al., 2005).

Although we often rely on in vivo data for informing TK, in vitro tools provide a relatively abundant and useful source of information (Donato et al., 2008). While these methods have long been used to assess expression of drug metabolizing enzymes in treated and untreated,

primary and immortalized cells in a more limited case-by-case basis (Geng and Strobel, 1995; Raunio et al., 1999; Swanson, 2004), omics technologies can be applied to broadly assess metabolic capacity between cell types of normal and abnormal phenotypes (Vondracek et al., 2001, 2002; Hedberg et al., 2001; Staab et al., 2008). Recently, an in vitro model of buccal epithelial tissue was used to examine the expression of carbonyl metabolizing enzymes in normal human basal and differentiated keratinocytes, as well as in immortalized malignant human keratinocytes (Cedar et al., 2007; Staab et al., 2008). Such approaches can inform the metabolic capacity of cells at a given stage of development (e.g., proliferation vs. differentiation) and, perhaps, the differential metabolic capacities of normal, pre-malignant, and malignant cells.

2.3.2.6. *Toxicogenomic Activities at the U.S. Food and Drug Administration (U.S. FDA)*

The U.S. Food and Drug Administration (U.S. FDA) initiated incorporating genomic data into their drug evaluation process, and thus, is a leader in this regard. It began to incorporate toxicogenomics data into their assessment and regulatory decisions following the voluntary submission of data by the industry for screening of drugs. Furthermore, the U.S. FDA has developed a draft guidance document to cover industry's submission of pharmacogenomic data (U.S. FDA, 2003). This guidance furthers scientific progress in the field of pharmacogenomics and facilitates the use of pharmacogenomic data in informing regulatory decisions. The draft guidance encourages, but again does not require, voluntary submission of microarray data from exploratory studies. This guidance does not include use of genetic or genomic techniques for the purposes of biological product characterization or quality control (e.g., cell bank characterization, bioassays). It also does not refer to data resulting from proteomic or metabolomic techniques. In addition, minimum information standards for microarray experiments for publications and submission to public repositories have been developed (Ball et al., 2004; Brazma et al., 2001).

The MicroArray Quality Control (MAQC) Consortium is a scientific community-wide effort, spearheaded by U.S. FDA scientists. The MAQC effort was developed to bring researchers from government, industry, and academia together to tackle issues of variability and contribute to the standardization of microarray procedures (Anonymous, 2006; Casciano and Woodcock, 2006; Frueh, 2006; Dix et al., 2006; Ji and Davis, 2006; Canales et al., 2006; Shippy et al., 2006; Tong et al., 2006; Patterson et al., 2006; MAQC Consortium et al., 2006; Guo et al.,

2006). The two main objectives of the 1st phase of the MAQC (MAQCI) project are (1) to compare cross-platform and interlaboratory performance of currently available microarray technologies and (2) to identify potential sources of variability. Seven different microarray platforms (six commercially available platforms [Applied Biosystems, Affymetrix®, Agilent Technologies, GE Healthcare, Ilumina, and Eppendorf] and one private platform [the National Cancer Institute]) were tested by three independent laboratories. Each laboratory used five sample replicates derived from four titration pools of two highly characterized unique RNA samples. The working list of genes was refined to include 12,091 reference genes that were detected on each of the six high-density platforms. The MAQCI study demonstrates that there is good reproducibility within sites, between sites, and among the various platforms. These findings are promising for future incorporation of microarray data into risk assessment procedures (MAQC Consortium, 2006).

The performance of the microarray platforms was further evaluated in comparison to three distinct quantitative gene expression assays: Taqman, Standardized RT-PCR, and Quantigene. There was excellent correlation between microarray results and quantitative gene expression results. Several sources of limited incongruence were identified: a decreased sensitivity for low expression genes in the microarray platforms as compared to the gene expression technologies and some differences in probe location.

A toxicogenomic study in rats was used to validate the observed congruence of microarray platforms in a biologically relevant framework. Rat RNA samples were collected and processed following exposure to three chemicals (aristolochic acid, ridelline, or comfrey). Results from four of the microarray platforms indicated a high degree of conformity. gone findings was that gene lists generated using fold-induction criterion showed much greater concordance across platforms as compared to those generated by t-test *P* values alone, with the novel finding that comfrey exposure results in differential regulation of vitamin A, and copper in the liver of rats was detected across all platforms.

The MAQCI project observed high reproducibility of findings between different microarray platforms tested at multiple locations. Additionally, microarray results were well correlated with other available gene expression technologies. Consistent results were also acquired in the toxicogenomic study after exposing rats. These studies provide the stepping-stones for decreasing variability in microarray data and add standardized quality-control

measures. Taken together, the findings are an encouraging first step for the future incorporation of microarray data into risk assessments. While it is noted that these results were a comparison of the same sample in different laboratories, a future step may consider a comparison of samples prepared in-house in independent laboratories/institutions.

2.3.2.7. *Toxicogenomic Activities at U.S. Environmental Protection Agency (U.S. EPA)*

The U.S. EPA has also initiated the development of methods, research, and guidancefor using toxicogenomic data for a number of purposes including risk assessment (see U.S. EPA, 2002; U.S. EPA, 2003; U.S. EPA, 2004b; U.S. EPA, 2006b). This includes training U.S. EPA risk assessors in genomics (e.g., Risk Assessment Forum Genomics Training Courses), developing guidance and methodology documents (e.g., this project), and supporting numerous research activities that are expected to support chemical-specific risk assessment activities in the future.

As previously described, the U.S. EPA's SPC developed the Interim Policy on Genomics. This policy states "genomics may be used in U.S. EPA risk assessments on a case-by-case basis in a WOE [weight-of-evidence] approach" (U.S. EPA, 2002). Currently there is no U.S. EPA guidance for how to incorporate toxicogenomic data into chemical assessments. The Genomics Task Force produced a white paper *Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA* that identified four areas of oversight likely to be influenced by genomic data: the prioritization of contaminants and contaminated sites, environmental monitoring, reporting provisions, and risk assessment. The paper also identifies a critical need for (1) analysis and acceptance criteria for genomic information in scientific and regulatory applications, (2) methods for interpreting genomic information for risk assessment, and (3) determining a relationship between genomic changes and adverse outcomes (U.S. EPA, 2004b). In response to these needs, the Genomics Technical Framework and Training Workgroup of the SPC was established and has since developed an *Interim Guidance for Microarray-Based Assays: Data Submission, Quality, Analysis, Management, and Training Considerations* (U.S. EPA, 2006b). This guidance addresses genomic data submission, quality assurance, analysis, and management in the context of current possible applications by the U.S. EPA and the broader academic and industrial community. The guidance also identifies future actions that are envisioned to incorporate genomic information more fully into the U.S. EPA's

risk assessments and regulatory decision making (Dix et al., 2006). Furthermore, U.S. EPA has institutionalized a national center, the National Center for Computational Toxicology (NCCT; www.epa.gov/NCCT) with one of its goals being to analyze and understand the omics data using a systems biology approach. U.S. EPA has also initiated both internal and external discussion to strategize and recommend next steps in methods development for the use of genomic data in risk assessment. These activities include the Office of Research and Development's Computational Toxicology Workshop: *Research Framework, Partnerships and Program Development* (September, 2003; Kavlock et al., 2004) and the National Center for Environmental Assessment (NCEA) colloquium, entitled *Current Use and Future Needs of Genomics in Ecological and Human Health Risk Assessment* (U.S. EPA, 2003; http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=149984), both of which identify the need to perform a case study integrating toxicogenomic data in a chemical assessment. Further, NCCT conducted a 3-day Science Forum in May 2007, where over 400 scientists from the international community met to discuss issues relating to genomics and computational toxicology.

Currently, U.S. EPA has attempted to incorporate toxicogenomic data (mostly qualitatively) in hazard identification of a few environmental chemicals. Two U.S. EPA assessments, the cancer assessment for acetochlor and for dimethylarsinic acid, evaluated the available genomic data (U.S. EPA, 2004c; 2006d). In both cases, the toxicogenomic data informed the MOA.

Although U.S. EPA has evaluated toxicogenomic data during the course of risk assessments, it has not developed a formalized approach for the incorporation of these data into risk assessment. Therefore, case studies, when performed in an iterative, collaborative fashion, could reveal practical issues for developing approaches and needs in utilizing toxicogenomic data in risk assessment. A case study to assess how to evaluate and utilize genomic data in risk assessment can identify: risk assessment areas that genomic data can inform, criteria for toxicogenomic data inclusion, and approaches and methods for incorporating toxicogenomic data in risk assessments. Nevertheless, as the technology continues to advance, U.S. EPA must prepare for the future increase in genomic data availability and submission by identifying (1) areas of risk assessment where such data may be particularly useful, (2) acceptance criteria for inclusion of toxicogenomic data in risk assessment, (3) approaches for the use of

toxicogenomics in risk assessment, and (4) research needs for developing and designing future studies.

2.3.2.8. *Toxicogenomic Activities at Other Agencies and Institutions*

In addition to the U.S. FDA and U.S. EPA, a number of other federal agencies, nongovernmental organizations, nonprofit organizations, and industry have conducted several studies and are involved in various activities of toxicogenomics. The following is a selective list of activities in other agencies and institutions. It should be noted that the toxicogenomic activities are not limited to the following organizations.

In November of 2000, the National Institute for Environmental Health Sciences (NIEHS) Division of Extramural Research and Training (DERT) issued a request to participate in a national Toxicogenomics Research Consortium. The four goals were to

(1) enhance research in the broad area of environmental stress responses using microarray gene expression profiling;

(2) develop standards and practices that will allow analysis of gene-expression data across platforms and provide an understanding of intra and interlaboratory variation;

(3) contribute to the development of a robust relational database, combining toxicological endpoints with changes in gene expression profiles; and

(4) improve public health through better risk detection and earlier intervention in disease processes(http://www.niehs.nih.gov/research/supported/centers/trc/).

The outcome of this consortium initiated areas that could have a major impact on risk assessment and public health.

In November of 2003, the International Programme on Chemical Safety (IPCS) conducted a workshop on *Toxicogenomics and the Risk Assessment of Chemicals for the Protection of Human Health*. The specific objectives of this workshop were to

- Establish a scientific forum for dialogue among experts;
- Share information about ongoing scientific activities using toxicogenomics at the national, regional, and international levels;
- Discuss the potential of toxicogenomics to improve the risk assessment process for the protection of health from environmental exposure to chemicals, understanding the MOA

of environmental toxicants, and the relevance and scope of gene-environment interactions;

- Identify the near-term needs and necessary steps for enhancing international cooperation in toxicogenomics research for improving chemical safety; and

- Identify and discuss data gaps, issues, and challenges that may present obstacles to the use of toxicogenomics for the protection of human health from environmental exposures.

The IPCS Workshop was successful in achieving its objectives as a number of areas of common interest were identified. The Workshop also confirmed the widely held view that toxicogenomics has the potential to improve the specificity and range of methods used to predict chemical hazards and to inform and to help overcome a number of uncertainties involved in chemical-related risk assessment.

The International Life Science Institute's (ILSI) Health Environmental Science Institute (HESI) has several completed and ongoing activities on the use of toxicogenomics in risk assessment. In 2004, Environmental Health Perspectives published a mini monograph, Pennie et al. (2004), with several articles relating to use and application of toxicogenomic data and their implications to risk assessment. In addition, ILSI/HESI has undertaken a major and ongoing effort to develop a toxicogenomic database (http://www.hesiglobal.org/Committees/TechnicalCommittees/Genomics/EBI+Toxicogenomics.htm). Furthermore, ILSI has conducted workshops and training courses on the use of toxicogenomic data in risk assessment. In addition, there is a recent source of information and training material that is published as an NRC report (NRC, 2007a).

2.3.3. Current Challenges and Limitations of Toxicogenomic Technologies

One of the major challenges in using microarray data is its interpretation in particular, the functional interpretation of genomic data or linking alterations in gene expression to conventional toxicological endpoints, sometimes referred to as "phenotypic anchoring" poses several obstacles that must be overcome. Another issue is reproducibility/variability (Moggs, 2005; Hamadeh et al., 2002a, b) in risk assessment; however, the MAQCI project results demonstrate good reproducibility when using the same biological sample and platform.

Although genomic data likely will impact multiple areas of science, medicine, law, and policy in the near future, there are a number of applications where genomic data have already been used in decision-making process (e.g., biomarkers of disease in medicine). Nevertheless, there are a number of technical and analytical methodological hurdles that must be addressed before genomic data can play a role in regulatory decision-making. These limitations include the paucity of toxicogenomic data for chemicals due to the cost, technical difficulties of conducting the experiments, and data analysis (Shi et al., 2004; Smith, 2001). Evaluation of methodologies including both the technologies themselves as well as the data analysis methods also needs validation. Until gene expression changes can be definitely linked with adverse outcomes, it is likely that gene expression data will continue to be used in conjunction with other traditional toxicological endpoints. To resolve these issues, an iterative and collaborative research process between risk assessors and research scientists would be very beneficial.

Despite these shortcomings, toxicogenomic technologies and data can facilitate risk assessment in several ways: (1) evaluating biological pathways/MOA for a given chemical or class of chemicals; (2) replacing standard toxicity screening assays in regulatory batteries; (3) assessing characteristics of the dose-response relationship, especially extrapolating from high experimental doses to environmentally relevant concentrations; (4) understanding the variability of responses in different species, or in different organs or tissues; and (5) evaluating individual variability and individual susceptibility based on the different gene expression patterns, especially polymorphic genes.

2.4. CASE STUDY

2.4.1. Project Team

The methods development and case study project were performed collaboratively between the U.S. EPA and outside partners. Team members include U.S. EPA scientists at NCEA, the National Health and Environmental Effects Laboratory, the Integrated Risk Information System (IRIS), and regional offices, as well as outside partners at the NIEHS, the Hamner Institute for Health Sciences, and the U.S. EPA Science to Achieve Results (STAR) Bioinformatics Center at Rutgers and University of Medicine and Dentistry of New Jersey (UMDNJ). The team was multidisciplinary, including experts in developmental and reproductive toxicology, human health risk assessment, toxicogenomic data study design, and

toxicogenomic data analysis. The multidisciplinary team included expertise in male reproductive and developmental toxicology and toxicogenomics.

2.4.2. Chemical Selection

We conducted a literature review to identify candidate chemicals for the case study. The literature review focused on endocrine disrupting chemicals (EDCs) because of the expertise of the team members and the availability of microarray studies for a number of EDCs. The androgen-mediated male reproductive development toxicity pathway was identified as the best choice for the case study (Figure 2-1) for four reasons:

(1) Androgens are essential for a number of male developmental events and are required during gestation for the normal development of the male genital tract and sexual differentiation; thus, this toxicity pathway has relevance to in vivo outcomes;

(2) There are published studies for chemicals that affect androgen action (i.e., androgen antagonists and agonists) that support a relatively strong linkage between the MOA and the resulting toxicological outcome after exposure;

(3) There are some published toxicogenomic data, as well as ongoing research, on some of the EDCs that affect androgen action; and

(4) There are recent or ongoing U.S. EPA assessments for some of chemicals that affect androgen action.

2.4.2.1. *Six Candidate Chemicals*

Six candidate chemicals were identified and considered for the case study: linuron, procymidone, vinclozolin, di-(2-ethylhexyl) phthalate (DEHP), DBP, and prochloraz. The criteria for selecting a chemical for the case study were

- Relative abundance of available toxicogenomic data (preferably published data);
- Consistency of the toxicogenomic data set findings, as one indicator of high quality studies;
- Recent or ongoing U.S. EPA assessment; and
- Interest by U.S. EPA Program and/or Regional Offices in performing a case study on this chemical.

We developed criteria to evaluate these six chemicals (Table 2-1). We gathered information on the criteria by reviewing the toxicogenomic literature and about the status of each chemical's U.S. EPA human health risk assessment. The summary of the information presented in the table and text is limited as it reflects the information available at the time of the decision about the case study chemical (July 2005).

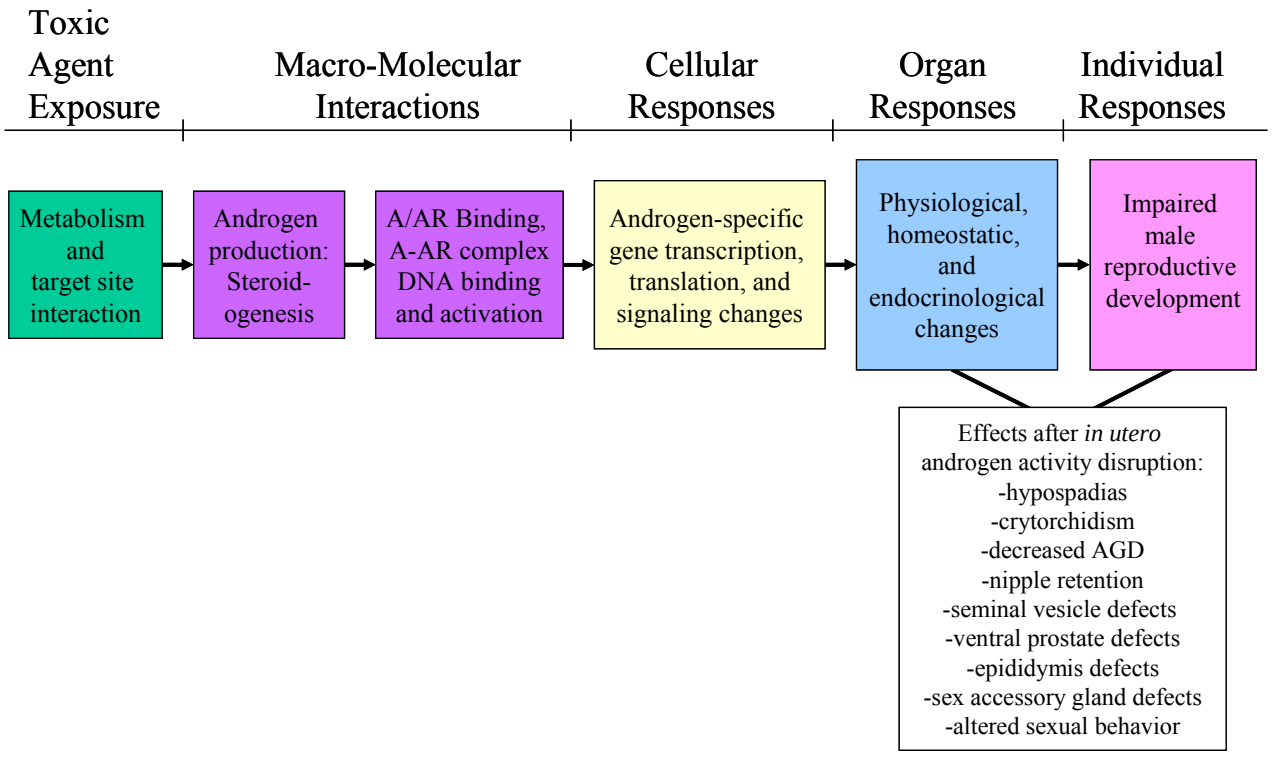

Figure 2-1. Androgen-mediated male reproductive development toxicity pathway.

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

Chemical	MOA(s)	U.S. EPA assessments (dates)?	Published TgX data (amount)?	Ongoing TgX studies?
Linuron	AR antagonist	IRIS Oral RfD, 1990; IRIS Cancer, 1993; OPP RED, 1995; OPP tolerance reassessment, 1999	Yes (low)	Ongoing
Procymidone	AR antagonist	Discussed in vinclozolin and iprodione OPP REDs; OPP tolerance reassessment, 1999	Yes (low)	Proposed (Gray, LE Jr., personal communication)
Vinclozolin	AR antagonist	OPP RED, 2000; 2002 OPP Final Risk Assessment; IRIS Oral RfD, 1992	Yes (low)	Yes
DEHP	Fetal testicular steroidogenesis inhibitor	Ongoing (IRIS)	Yes (high)	Yes
DBP	Fetal testicular steroidogenesis inhibitor	Ongoing; Internal review complete (IRIS)	Yes (high)	Yes
Prochloraz	Steroidogenesis inhibitor and AR antagonist	IRIS Oral RfD, 1989; IRIS Cancer, 1997	Yes (medium), but few studies focused on male repro tissues and/or endpoints	Proposed (Gray, LE Jr., personal communication)

AR, androgen receptor; OPP, Office of Pesticide Programs; RED, Reregistration Eligibility Decisions; RfD, reference dose; TgX, toxicogenomic.

[1]The information in this table reflects the available information at the time of the decision (July 2005).

2.4.2.2. *DBP Selected as Case Study Chemical*

All the candidate chemicals—except prochloraz—meet three of the four criteria for chemical selection: (1) a relative abundance of available toxicogenomic data, (2) a relatively consistent toxicogenomic data set, and (3) a recent (<5 years) or ongoing U.S. EPA assessment. Assessment of the 4th criteria was more subjective in nature, as individuals' opinions were queried. However, none of the five remaining chemicals were considered a poor choice. After discussion of the relative merits of each of the five chemicals, we selected DBP for the case study for the following reasons:

1) **Quantity and Quality of Toxicogenomic Data Set:**
   DBP and DEHP both have a relatively large and high-quality (based on consistency of findings) toxicogenomic data set. The DBP data set includes gene expression changes in genes known to be involved in the androgen-mediated male reproductive toxicity pathway, providing phenotypic anchoring to a number of the male reproductive developmental effects following high dose DBP in utero exposure. Additionally, there is one dose-response RT-PCR study using low-to-high in utero DBP doses that observed alterations in nine genes involved in steroidogenesis as well as other pathways (Lehmann et al., 2004).

2) **Application to Risk Assessment:**
   The DBP assessment may allow the case study to address some interesting questions that may have broad application to the use of toxicogenomics in risk assessment. These questions include

   - Do the toxicogenomic data provide information about multiple and/or additional MOA(s) for DBP?

   - Could toxicogenomic data be used to determine the adverse level for the reduction in fetal testosterone (T), the MOA for a large number of the male reproductive developmental endpoints after in utero DBP exposure?

3) **Availability of Draft Assessment:**
   At the time of chemical selection for this case study, the external review draft of the IRIS DBP assessment was being developed and, thus, available for use as a starting point for the case study. Risk assessment documents for the other candidate chemicals were either >5 years old, running the risk of needing more information incorporated for the case study, or too early in the stage of the process to utilize a draft document.

2.4.3. Case Study Scope

After DBP was selected for the case study, the scope of the case study was further defined. The DBP case study is limited to effects on male reproductive development because (1) these endpoints are the current focus in the IRIS assessment as they occur in the lower dose range; (2) the team members have expertise in reproductive and developmental biology and toxicology; and (3) some of these endpoints have been associated with a number of the gene and pathway alterations, thus providing a phenotypic anchor. After reviewing the data sets for DBP (see Chapter 3), the initial focus on androgen-mediated male reproductive developmental effects (see 2.3.2) was broadened to include all male reproductive developmental effects, and not just those affecting androgen action, because DBP affects the other pathways (e.g., *Insl3*) as well as the androgen pathway.

The approach design used a health assessment model, focusing on utilizing genomic data to inform the hazard characterization and dose-response steps of risk assessment. Thus, exposure assessment step was not included in this approach. While there are many successes and ongoing efforts utilizing genomics in exposure assessment, both in ecological and human health risk assessment, but these will not be covered in this document.

The DBP case study, focuses on considering the various types of information useful to hazard characterization and dose-response that the genomic data may inform. The incorporation of toxicogenomic data into risk assessment includes both a quantitative and qualitative use of these data. However, the DBP case study is limited to the use of genomic data to inform the qualitative aspects of risk assessment because of the lack of available dose-response toxicogenomic data for DBP. The application of toxicogenomic data to quantitative aspects, such as TK modeling and dose-response assessment, is discussed in this document (see Chapters 3 and 7). This general discussion includes considerations that may be useful to a risk assessor evaluating genomic data.

3. DBP CASE STUDY APPROACH AND EXERCISE

This chapter presents a description of the approach to evaluating toxicogenomic data in risk assessment, and it also describes the first three steps of the DBP case study. Our strategy for evaluating genomic data for risk assessment was to design a flexible and systematic approach that would provide a thorough evaluation of the genomic data set for a particular chemical, while still accommodating different risk assessment practices. The discussion includes both (1) generic considerations for evaluating the data set for any chemical; and (2) explanations of how these issues were considered for the DBP case study.

3.1. EVALUATING DBP IRIS ASSESSMENT EXTERNAL REVIEW DRAFT

The case study approach begins with an evaluation of the existing DBP external review draft IRIS assessment document (see Figure 3-1). Use of the ongoing IRIS DBP assessment external review draft as the starting point allowed us to take advantage of (1) the compilation of the toxicity and human data sets, allowing us to focus on the toxicogenomic data set evaluation (2) data gaps that were identified, thus, providing possible questions that the toxicogenomic data may be able to address.

The IRIS Assessment for DBP was in progress when this toxicogenomic case study on DBP was initiated (2005). The IRIS Agency Review had been completed, and the Toxicological Review and IRIS Summary were in Interagency Review. Upon completion of the Interagency Review, the Toxicological Review and IRIS Summary were released for public comment in mid-July 2006. The Peer Review Panel meeting was held July 28, 2006 (http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=155707).

There are extensive studies documenting developmental toxicity of dibutyl and the metabolite, monobutyl phthalate, in rodents (Barlow et al., 2004; Barlow and Foster, 2003; Mylchreest et al., 2002, 2000, 1999, 1998; Ema and Miyawaki, 2001a, b; Ema et al., 2000a, b 1998, 1997, 1996, 1995, 1994, 1993; See Chapter 4 for further details). DBP exposure during a critical window of development in late gestation to the developing male rat fetus causes a variety of malformations of the reproductive tract structures. These include hypospadias; decrease in

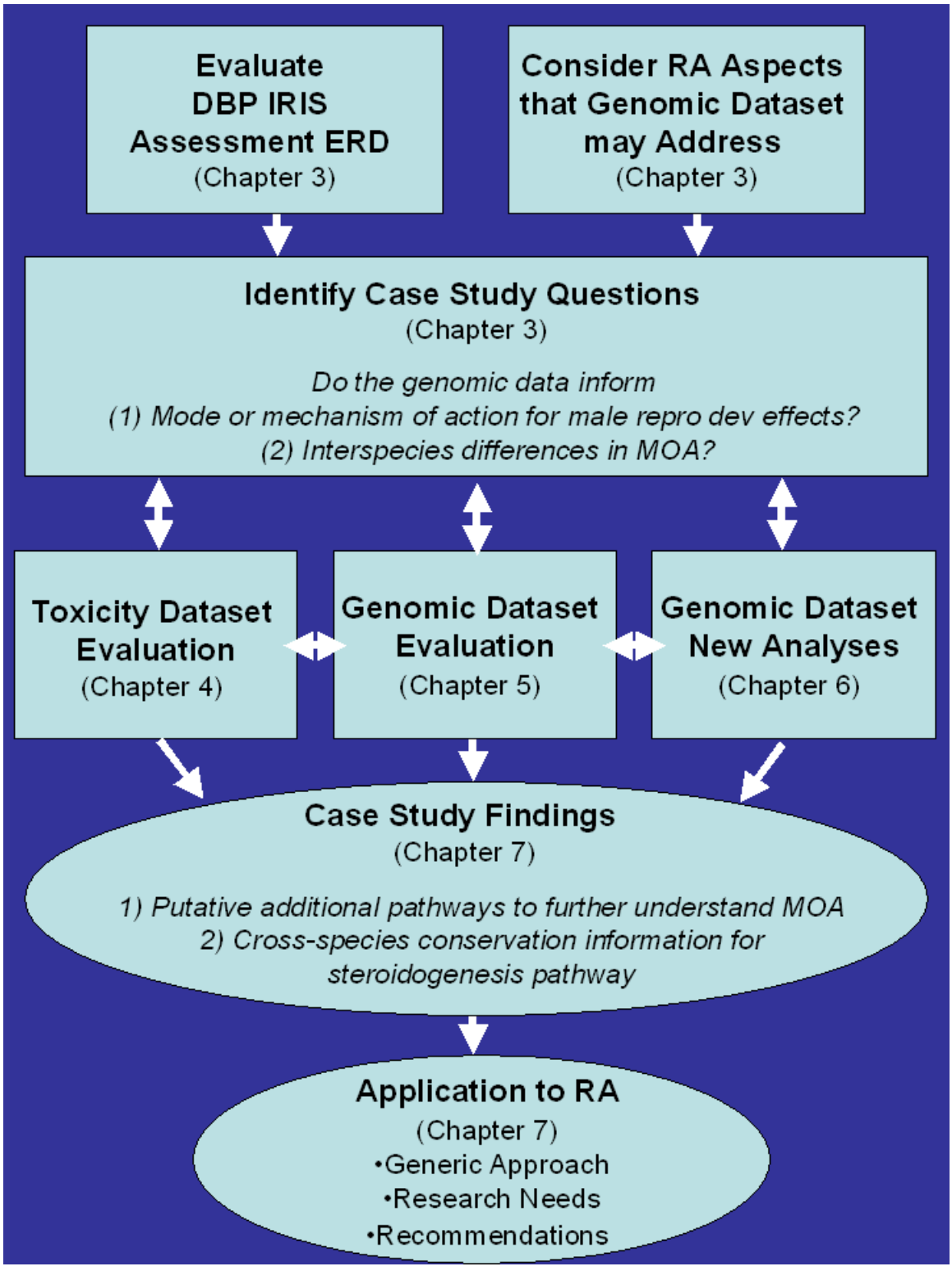

Figure 3-1. DBP case study approach for evaluating toxicogenomic data for use in health assessment. Evaluation steps in the case study process are shown in rectangles. Findings or products of the case study are shown in ovals. ERD, external review draft. Numbers in parentheses indicate report chapters where the case study step is described.

anogenital distance (AGD); delayed preputial separation (PPS); agenesis of the prostate, epididymis, and vas deferens; degeneration of the seminiferous epithelium; interstitial cell hyperplasia of the testis; and retention of thoracic areolas and/or nipples (Bowman et al., 2005; Kleymenova et al., 2005a; Barlow et al., 2004; Kim et al., 2004b; Barlow and Foster, 2003; Fisher et al., 2003; Higuchi et al., 2003; Mylchreest et al., 2002, 2000, 1999, 1998; Ema et al., 2000b, 1998, 1997, 1994; Saillenfait et al., 1998). For example, Mylchreest et al. (2000) observed retained areolas and/or nipples after exposure to 100 mg/kg-d DBP and observed a no-effect level at 50 mg/kg-d.

Figure 3-2 shows the studies that were candidates for the development of the reference dose (RfD) in the IRIS DBP external review draft assessment (U.S. EPA, 2006a). The point of departure (POD) selected for derivation of the RfD for all exposure durations (acute, short-term, subchronic, and chronic) is the no-observed-adverse-effect level (NOAEL) of 30 mg/kg-d for reduced fetal testicular T (Lehmann et al., 2004). In this study, a statistically significant decrease in T concentration in the fetal testis was detected at 50 mg/kg-d. The reduction in fetal testicular T is a well characterized MOA that occurs after in utero DBP exposure during the critical window and initiates the cascade of events for a number of malformations in the developing male reproductive tract. Studies using RT-PCR, immunochemical staining, and radioimmunoassay for T levels showed a decrease in protein and mRNA for several enzymes in the biochemical pathways for cholesterol metabolism, cholesterol transport, and for T biosynthesis (also called steroidogenesis more generally) in the fetus (Plummer et al., 2005; Thompson et al., 2005, 2004; Lehmann et al., 2004; Barlow et al., 2003; Fisher et al., 2003; Shultz et al., 2001). Collectively, these studies document that exposure to DBP disrupts steroid synthesis in the fetal testis. Thompson et al. (2004) established that following in utero exposure to 500 mg/kg-d, the T levels in the testes return to normal after the metabolites of DBP are cleared from the circulation. The malformations induced by exposure to 500 mg/kg-d persist into adulthood (Barlow et al., 2004; Barlow and Foster, 2003). Thus, although the inhibition of T synthesis is reversible, the biological effects resulting from the decrease in T during the critical developmental window are irreversible.

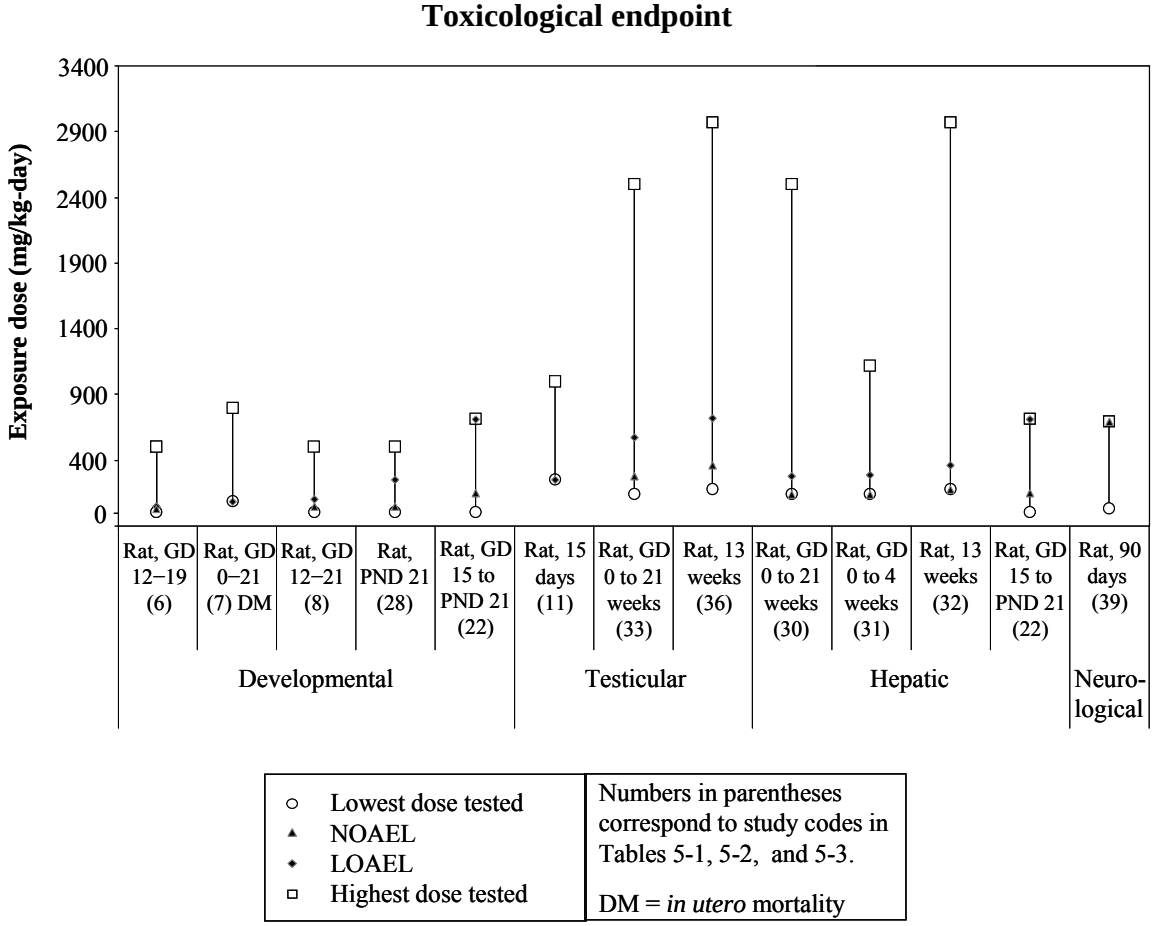

Figure 3-2. Exposure response array for candidate endpoints for the point of departure (POD) in the IRIS DBP assessment external review draft. The studies are arrayed by toxicological endpoint. Within each toxicological endpoint, the studies are arrayed by duration of exposure, shortest to longest. DM is in utero mortality. The open circle is the lowest dose tested, and the filled triangle is the NOAEL (no-observed-adverse-effect level, the filled diamond is the LOAEL (lowest-observed-adverse-effect level), and the open square is the highest dose tested. The numbers in parentheses refer to study numbers in tables in the external review draft of Toxicological Review of Dibutyl Phthalate (U.S. EPA, 2006a) and are as follows: 6, Lehmann et al. (2004); 7, 30, 31, 32, 33, and 36, National Toxicology Program (NTP, 1995); 8, Mylchreest et al. (2000); 11, Srivastava et al. (1992); 22, Lee et al. (2004); 28, Zhang et al. (2004); and 39, BASF (1992). GD, gestation day; PND, postnatal day.

3.2. CONSIDERATION OF RISK ASSESSMENT ASPECTS THAT TOXICOGENOMIC DATA MAY ADDRESS

While microarray and RT-PCR data have been used to inform the MOA of a chemical, appropriate genomic data have the potential to inform TK, dose-response, interspecies and

intraspecies differences in TK or TD, and be utilized as biomarkers of exposure or effect (see Figure 3-3). We considered the use of toxicogenomic data in health assessments and the many types of information useful to hazard characterization, dose-response analysis, and risk characterization. Toxicogenomic data have been successful in providing information about the molecular events altered in the mechanism of action, and in some cases, information about TD or TK MOA events, intra- and interspecies differences in molecular responses (see Figure 3-4).

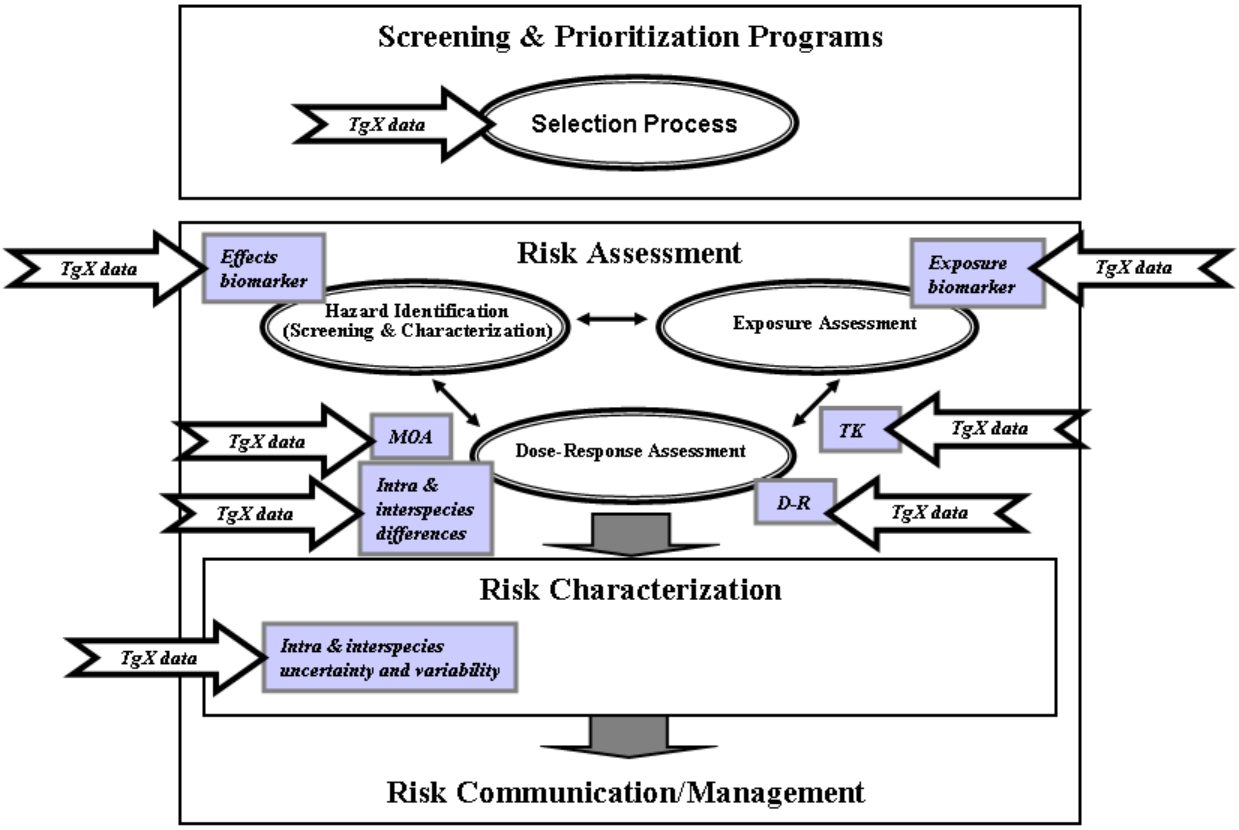

Figure 3-3. Potential uses of toxicogenomic data in chemical screening and risk assessment. Genomic data from appropriately designed studies have the ability to inform multiple types of information and in turn, steps in screening and risk assessment. Arrows with "TgX data" (toxicogenomics data) indicate the types of information these data can provide. Shaded boxes indicate some of the types of information that are useful in risk assessment.

In this case study, chemical screening and exposure assessment were not considered. We considered the use of toxicogenomic data in health assessments and the many types of information useful to hazard characterization, dose-response, and risk characterization. Toxicogenomic data have been successful in providing information about the molecular events altered in the mechanism of action, and, at times, can provide information about the TD or TK key events of the MOA (see Figure 3-4). Data from appropriately designed toxicogenomic studies could be used to inform intraspecies and interspecies differences in molecular responses.

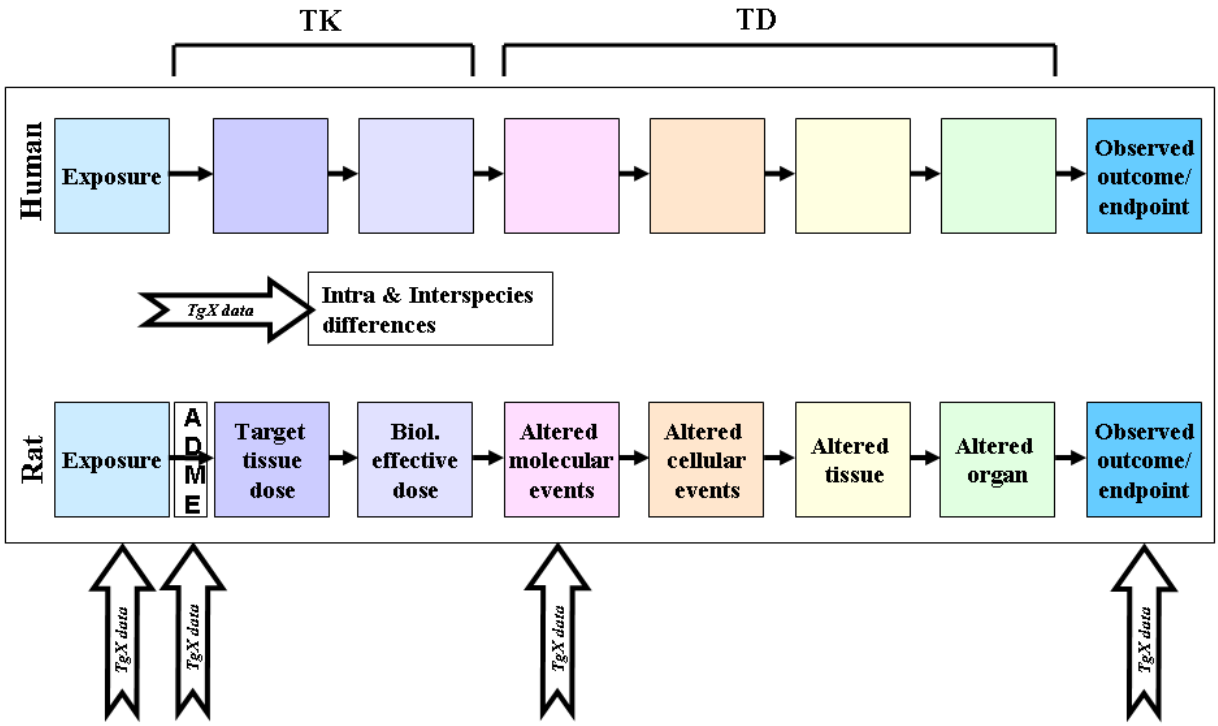

Figure 3-4. Potential uses of toxicogenomic data in understanding mechanism of action. The process from exposure to outcome encompasses all of the steps of the mechanism of action, including both toxicokinetic (TK) and toxicodynamic (TD) steps. Available toxicogenomic (TgX) data, such as microarray data and other gene expression data, can provide information about altered molecular events, at the gene expression level. In turn, TgX data can be used to inform intraspecies and interspecies differences in molecular responses. Appropriate TgX data could also inform internal dose and intra- and interspecies differences in internal dose. ADME, absorption, distribution, metabolism, and excretion.

3.2.1. Informing Toxicokinetics

Characterizing the absorption, distribution, metabolism, and excretion (ADME) of environmental toxicants is important for both the understanding and application of MOA information in predicting toxicity in health risk assessments. Differences in TK across species, individuals, and exposure patterns (routes, level, duration, and frequency) can lead to different biological effects for the same total amount of exposure to a chemical. It is well established (U.S. EPA, 2006d) that a quantitative understanding of chemical TK (e.g., using PBPK models) can be useful in analyzing dose-response data and extrapolating across species, individuals, and exposure patterns. The principles of these uses for TK are the same, regardless of whether the endpoint(s) are in vivo toxicity endpoints (e.g., pup weight) or molecular precursor events (e.g., toxicogenomic changes), and will not be reviewed here. However, the inverse question—how toxicogenomic data can inform TK—has not been fully explored. Here we consider whether toxicogenomic data could be useful for understanding four aspects of a chemical's TK: (1) identification of potential metabolic and clearance pathways; (2) selection of an appropriate dose metric; (3) inter and intraspecies differences in metabolism; and (4) TK/TD linkages and feedback. Each of these applications is discussed below. Finally, the available toxicogenomic data for DBP are evaluated for use in informing TK.

3.2.1.1. *Identification of Potential Metabolic and Clearance Pathways*

While TK studies, themselves, are designed to help understand the pathways for metabolism and clearance of xenobiotics, toxicogenomic data may provide important complementary information as to what enzymes and tissues may be involved in metabolism. For example, many xenobiotics induce the expression of the Phase I and II enzymes that are responsible for their clearance. Thus, toxicogenomic data showing expression changes in genes such as cytochrome P450s in a particular tissue may implicate their involvement in metabolizing the compound. While such toxicogenomic data may confirm the major sources of metabolism or clearance, they may also identify minor TK pathways relevant for inducing toxicity. However, toxicogenomic changes alone are insufficient to conclude that there is a corresponding increase in a protein or activity, or is relevant to the ADME of the chemical of interest. Ultimately,

toxicogenomic data may be most useful for generating hypotheses about metabolism and clearance pathways that can be tested with additional TK studies.

3.2.1.2. *Selection of Appropriate Dose Metrics*

Due to inherent differences in TK across species, individuals, and exposure patterns, dose-response relationships are best established based on an internal measure of a biologically effective dose as opposed to an external or applied dose. However, an understanding of TK alone may provide a multitude of different options for this internal "dose metric," such as blood or tissue concentrations of the parent or metabolites, or rates of formation of reactive compounds. Thus, a key question in utilizing TK data for dose-response analyses and extrapolation is dose metric selection, which depends on the determination of the active chemical species and the MOA of toxicity. There often may be more than one biologically plausible choice of dose metric, which contributes to the uncertainty in the dose-response analysis. The potential utility of toxicogenomic data is that they are intended to represent earlier biological effects, and, thus, are closer both spatially and temporally with the interaction between the active chemical species and endogenous cellular molecules than more readily observable outcomes. Thus, toxicogenomic data can, in principle, provide biological support for the choice dose metric. Different predictions for internal dose can be statistically analyzed along with toxicogenomic changes that inform TD to determine the dose metric that is best correlated.

3.2.1.3. *Intra- and Interspecies Differences in Metabolism*

Perhaps the most straightforward application of toxicogenomic data in TK analysis is to characterize intra- and interspecies differences in metabolism. Data from polymorphisms is one type of genomic data that can be extremely useful to informing intraspecies differences. Across species, data on differential expression of different isozymes genes may be indicative of differences in overall metabolizing capacity and affinity. In addition, toxicogenomic data may be informative as to whether the tissue distribution of metabolizing enzymes may be different across species. Within species, interindividual variability in metabolizing capacity and/or affinity due to differences in expression or genetic polymorphism can greatly influence the overall TK of a chemical. For example, genetic polymorphisms in aldehyde dehydrogenase-2 (*Aldh2*) can result in an increase in blood acetaldehyde levels following alcohol consumption,

thereby leading to overt health effects (see Ginsberg et al., 2002). Similarly, data on CNPs can provide information (Buckley et al., 2005) with direct impact on TK. For example, some individuals possess different copy numbers of *Cyp2d6* that influence their response to pharmaceuticals (Bodin et al., 2005). When the impacts of gene expression levels and polymorphisms on enzyme levels and function are known (i.e., preferably confirmed by measurement of enzyme level), this information can either be used to characterize the difference in a predicted dose metric for a subpopulation relative to the most common alleles, or it can be used in probabilistic (e.g., Monte Carlo) analyses to characterize the impact of population variability.

3.2.1.4. *Toxicokinetic/Toxicodynamic (TK/TD) Linkages and Feedback*

Ultimately, toxicogenomic data may provide a crucial element for linking together TK and TD models into more comprehensive biologically based dose-response (BBDR) models (Daston, 2007). With an appropriate dose metric, one can link the TK predictions for a chemical (e.g., tissue concentration of a metabolite) with toxicogenomic changes (e.g., change in mRNA transcript level) that, in turn, are linked through a TD model to alterations in cellular constituents and, ultimately, frank effects. Furthermore, toxicogenomic data may be useful in providing the link by which the TD feedback of gene and protein expression changes on TK (e.g., such as enzyme induction) can be modeled.

3.2.1.5. *Research Needs for Toxicogenomic Studies to Inform Toxicokinetics*

Changes in gene expression can be highly labile and vary as a function of dose and time. Thus, identification of appropriate dose metrics involves detection of relevant gene changes as well as the moiety that caused the changes. Therefore, simultaneous data collection of toxicogenomic data and tissue concentrations of the relevant chemical species would be beneficial. Concerning interspecies extrapolation, it is important to mine toxicogenomic data for potential indicators of species differences in metabolism. For intraspecies variability, it is important to assess the potential impact of polymorphisms in Phase I and II enzymes. Microarray data may also be useful for identifying life stage and gender differences in relative expression of enzymes involved in the TK of the chemical of interest.

3.2.1.6. *DBP Case Study: Do the Available Toxicogenomic Data Inform TK?*

We considered whether the available toxicogenomic data set informs TK. A greater level of detail is presented for TK here (Chapter 3) than for MOA because the latter subject is considered in greater detail in the subsequent chapters. This section provides examples of considerations that may be helpful to risk assessors examining whether the available toxicogenomic data can inform TK for their chemical of interest.

The TK of DBP is reviewed in U.S. EPA (2006a) and is summarized briefly here for context. Following ingestion, DBP is primarily hydrolyzed to monobutylphthalate (MBP) in the gastrointestinal tract and enters systemic circulation through the portal blood. MBP undergoes glucuronidation in the liver, and both free and glucuronidated MBP circulate in serum and are subsequently excreted in urine. While there are a number of TK studies in rats, little such data are available in humans, particularly for known exposures to DBP. The available data suggest that free MBP is responsible for the effects on T biosynthesis. In terms of TK pathways, the data set did not lead to the identification of alternative metabolic pathways for DBP.

Toxicogenomic data could inform dose metric selection in two broad ways: relating the metabolite to the gene expression or using gene expression as the dose metric. In a more traditional approach, expression changes in genes of interest can be related to a chemical moiety in a target tissue of relevance (or convenience). For example, Lehmann et al. (2004) provides a dose-response analysis of gene expression following DBP exposure. However, this study is of limited value for extrapolation without TK data (e.g., tissue concentrations of MBP). Ideally, TK data could be collected at various time points following various doses, but this would require a large number of fetuses. In the absence of such empirical data, analyses could be performed using physiologically based TK modeling, but none have yet been attempted. Such an approach might utilize TK studies for DBP and attempt to reconstruct the exposure scenarios in the toxicogenomic studies with the intent to predict the MBP concentration in a target tissue (or blood) at the time points where toxicogenomic samples were obtained.

A second and more complex approach might be to use a toxicogenomic change as a dosimeter (or "biomarker"), which may obviate the need for TK data altogether. For example, the microarray study of Wyde et al. (2005) reports changes in maternal liver *Cyp2b1*, *Cyp3a1*, and estrogen sulfotransferase mRNA levels following DBP exposure. Not only do these gene expression changes serve as potential biomarkers, but also suggest that there may be related

changes in metabolic biomarkers (i.e., metabonomics) because these enzymes have roles in lipid and hormone synthesis, in addition to xenobiotic metabolism. Although these changes may have no relationship to the toxic endpoint of interest, it may be possible to establish, for instance, that an increase in a specific maternal liver mRNA is correlated with a decrease in a specific mRNA in the fetal testis. Indeed, Wyde et al. (2005) show that maternal liver estrogen sulfotransferase gene expression increases in a dose-dependent manner from 10 to 500 mg/kg, while over nearly the same dose ranges, Lehmann et al. (2004) show a dose-dependent decrease in male fetal testicular *Scarb1*, *Star*, *Cyp11a1*, and *Cyp17a1* mRNA levels. Establishing such correlations in humans is not feasible; however, if similar correlations might be found in more accessible tissues. For example, if there were strong correlations between changes in rat maternal blood cell estrogen sulfotransferase mRNA and changes in a fetal testis mRNA of interest, then elevations in human blood cell estrogen sulfotransferase mRNA might be indicative of DBP-related changes in human male testis.

With respect to interspecies extrapolation and interindividual variability, the lack of adequate human TK data precludes quantitative extrapolation, a situation that cannot be remedied with toxicogenomic data (unless, as discussed above, a toxicogenomic-based dosimeter/biomarker is developed). For instance, available blood measurements of MBP in humans were taken from spot samples in the general population where the individual exposure patterns were unknown. Although differences were observed in the ratio of free to conjugated MBP in serum as compared to the rat, these data are insufficient for quantitative interspecies extrapolation because in order to replace administered dose as a dose metric, it is necessary to determine the *absolute*, not the *relative*, level of free MBP in serum as a function of exposure. The rat data also suggest that enzyme induction occurs as Wyde et al. (2005) provided toxicogenomic evidence that exposure to 50 and 500 mg/kg DBP leads to an increase in rat liver UDP glucuronsyltransferase 2B1 (*Ugt2b1*) mRNA levels. More TK analysis would be required to ascertain whether this induction in rats occurs at levels that are relevant to low-dose exposures. Moreover, this may indicate that such induction occurs in humans and that this response may increase interindividual sensitivity to DBP toxicity. With regard to human TK, none of the available toxicogenomic data on DBP provide any information on DBP interindividual TK variability such as polymorphisms in glucuronyltransferases responsible for metabolizing MBP. Finally, we considered the potential for TK/TD linkages with the available

data. It is also likely that in order for TK and toxicogenomic data to be integrated for quantitative dose-response analysis, more sophisticated BBDR models will need to be developed. Using such an approach, it may be feasible to relate changes in genes involved in T production to quantify testicular T levels (see Figure 3-5).

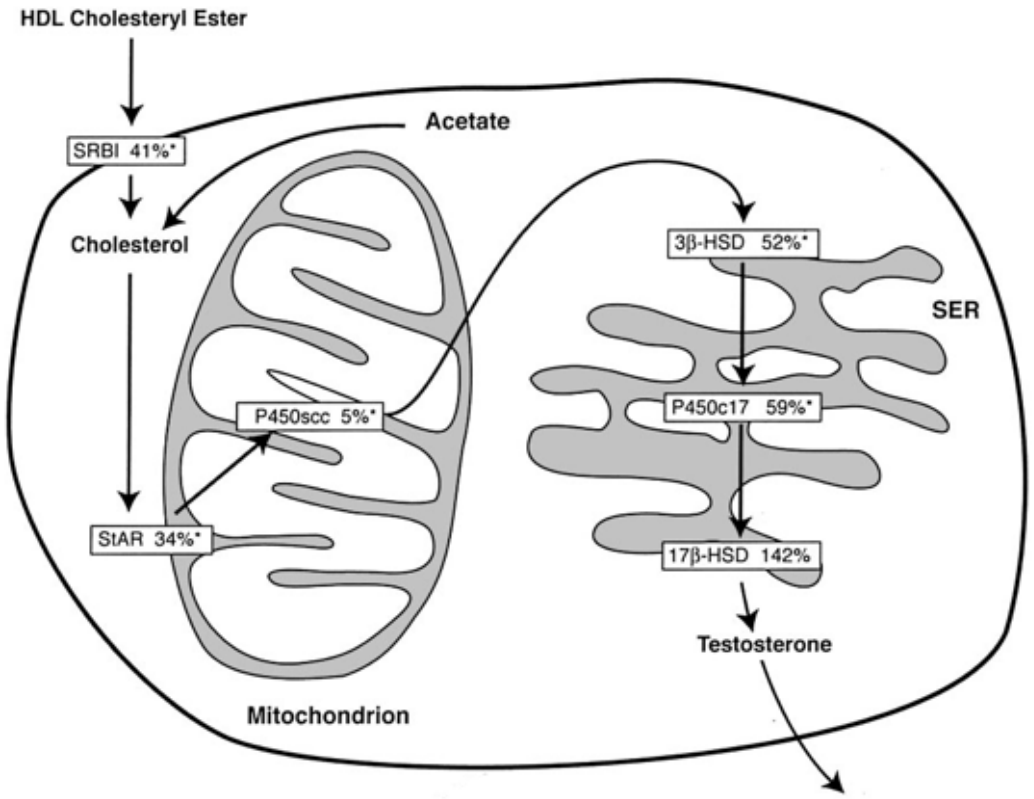

Figure 3-5. The fetal Leydig Cell in the fetal testis. The boxes represent genes involved in the biosynthesis of T; the percentages (%) represent % control gene expression in fetal testis of dams treated with 500 mg/kg-d DBP.

Source: Adapted from Barlow et al. (2003).

Briefly, the deleterious effects of DBP appear to be mediated by MBP, which causes a down regulation of cholesterol transporters across the cell membrane (SCARB1) and mitochondrial

inner membrane (StAR), as well as the down regulation of two enzymes involved in converting cholesterol to T, CYP11a1, and CYP17a1 (Liu et al., 2005; Lehmann et al., 2004; Barlow et al., 2003; Shultz et al., 2001). Thus, it may be possible to relate DBP and/or MBP levels to reductions in cholesterol transporter (e.g., SCARB1 and StAR) and biosynthetic (CYP11a1 and CYP17a1) mRNA, protein, and/or activity levels. Changes in these parameters may then be modeled to predict changes in testicular T levels, which may subsequently be correlated to developmental toxicities.

3.2.2. Informing Dose-Response

Toxicogenomic data that informs TK can be useful for informing or improving dose-response analysis because it may improve the prediction of the dose metric of selection among alternative dose metrics. However, use of toxicogenomic data as an endpoint in dose-response analysis has not been extensively explored. Some dose-response microarray studies relating gene ontology categorization of gene expression changes have utilizing BMD analysis to determine PODs as a function of dose (Thomas et al., 2007; Andersen et al., 2008).

3.2.3. DBP Case Study: Do the Toxicogenomic Data Inform Dose-Response?

The available toxicogenomic data set for DBP can be useful for dose-response analysis. Specifically, Lehmann et al. (2004) showed that fetal testicular testosterone was significantly reduced at 50 mg/kg-d or higher. A Western analysis of four proteins involved in testosterone synthesis indicated that two proteins were significantly decreased at 50 mg/kg-d, a third protein was also decreased at this dose, albeit insignificantly, while a fourth protein was only reduced at 500 mg/kg-d. It would be helpful to use proteomics analysis to assess protein expression on a global level. RT-PCR analysis findings confirmed that the mRNA of all four genes was significantly reduced starting at 50 mg/kg-d. Unfortunately, there are currently no dose response microarray studies to assess the global expression over a dose range. However, this one dose response gene expression study does support the role of steroidogenesis and cholesterol transport genes in the decreased in testicular testosterone after in utero DBP exposure.

3.2.4. Informing Toxicodynamics/Mechanism and Mode of Action

There are numerous examples where toxicogenomic data have been used to inform the mechanism or MOA for a chemical, and there are a small number of examples where such data have been used corroboratively for risk assessment decisions (see Chapter 2).

3.2.4.1. *General Considerations: Mechanism and Mode of Action*

One feature of the approach (Figure 3-1) is the evaluation of the toxicity and toxicogenomic data sets in conjunction. The purpose of the evaluation was to consider the relevance of gene expression changes with respect to specific endpoints of interest identified in the toxicity data set. In addition, using this approach could provide connections between affected pathways (toxicogenomic data set) and endpoints affected (toxicity data set), which may, in turn, inform modes or the mechanism of action, as illustrated by Figure 3-6. Chapter 2 and the glossary describe the distinction between the definitions for mechanism of action and MOA. By linking the pathway and MOAs identified in this approach, pathways may be matched with and inform the mechanism of action for a chemical.

The decision logic of the MOA framework in the U.S. EPA Cancer Guidelines (U.S. EPA, 2005) could be utilized in this step of the approach (i.e., the available data are considered in light of a hypothesized MOA and follow a decision tree). However, the approach outlined here is designed to specifically consider the genomic data for informing MOA which is different from the goal of the MOA Framework.

This approach is best suited to instances where comparable study designs between the toxicity/epidemiology and toxicogenomic data sets are available. For example, toxicogenomic and toxicity studies performed in the same species, using similar doses, similar exposure intervals, and assessing the same organ or tissue would be ideal for utilizing this approach. For the DBP toxicity (see Chapter 4) and toxicogenomic data sets (see Chapter 5), there is some comparability across some of the studies—i.e., some toxicity and toxicogenomic studies were performed at the same doses with similar exposure intervals, in the same species and strain, and assessed some of the same organs (e.g., testis). However, no two studies are comparable for all study-design aspects, such as precise timing of exposure and time of assessment.

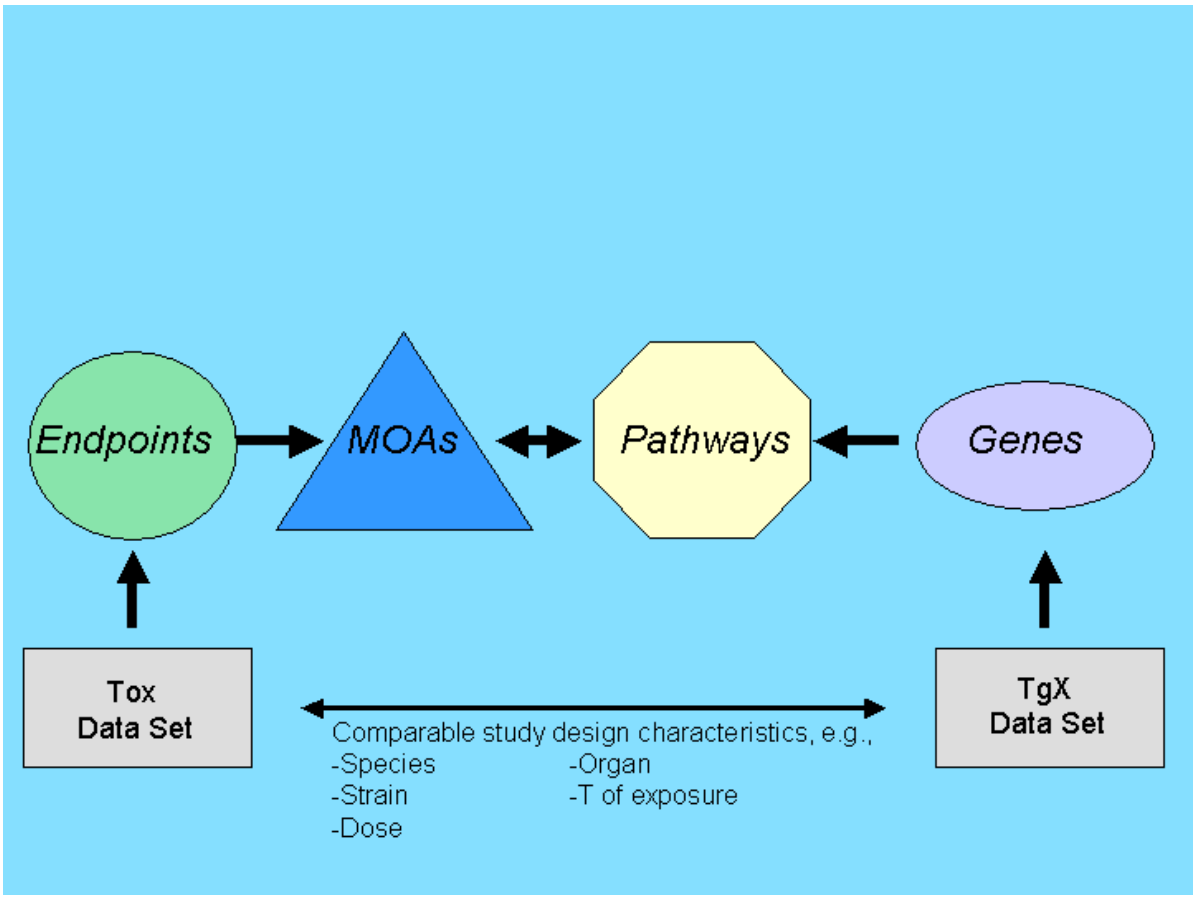

Figure 3-6. Approach to utilizing toxicity and toxicogenomic data for identifying affected pathways and candidate modes and mechanism of action. Toxicogenomic data can be analyzed for differentially expressed genes (DEGs) and, in turn, grouped into affected pathways. Toxicity data can provide information about affected endpoints. Toxicogenomic and toxicity data can inform mechanism of action, including MOAs, for a chemical by relating the endpoints and the pathways. Such an approach requires similar study parameters (e.g., dose, species, duration of exposure) for the toxicity and toxicogenomic studies. TgX, toxicogenomic.

3.2.4.2. *DBP Case Study: MOAs for Male Reproductive Developmental Effects*

Developmental toxicity studies (reviewed in Chapter 4) and toxicogenomic studies (reviewed in Chapter 5) have contributed to a good understanding of DBP as a chemical that has multiple MOAs. Two well characterized MOAs, a reduction in fetal testicular T, and a reduction in *Insl3* signaling activity explain a number of the observed male reproductive developmental abnormalities. Some other observed abnormalities are not explained by these two MOAs, suggesting that there are additional MOAs for DBP. Acknowledging that there are additional data not presented in Figure 3-7, this figure attempts to show where there is agreement in the scientific community (based on reproducibility of microarray and RT-PCR studies) about affected pathways and the well characterized MOAs for DBP. There are some endpoints and pathways that need further characterization and, as a result, we were interested in determining whether the toxicogenomic data could be used to associate the DBP MOAs and endpoints.

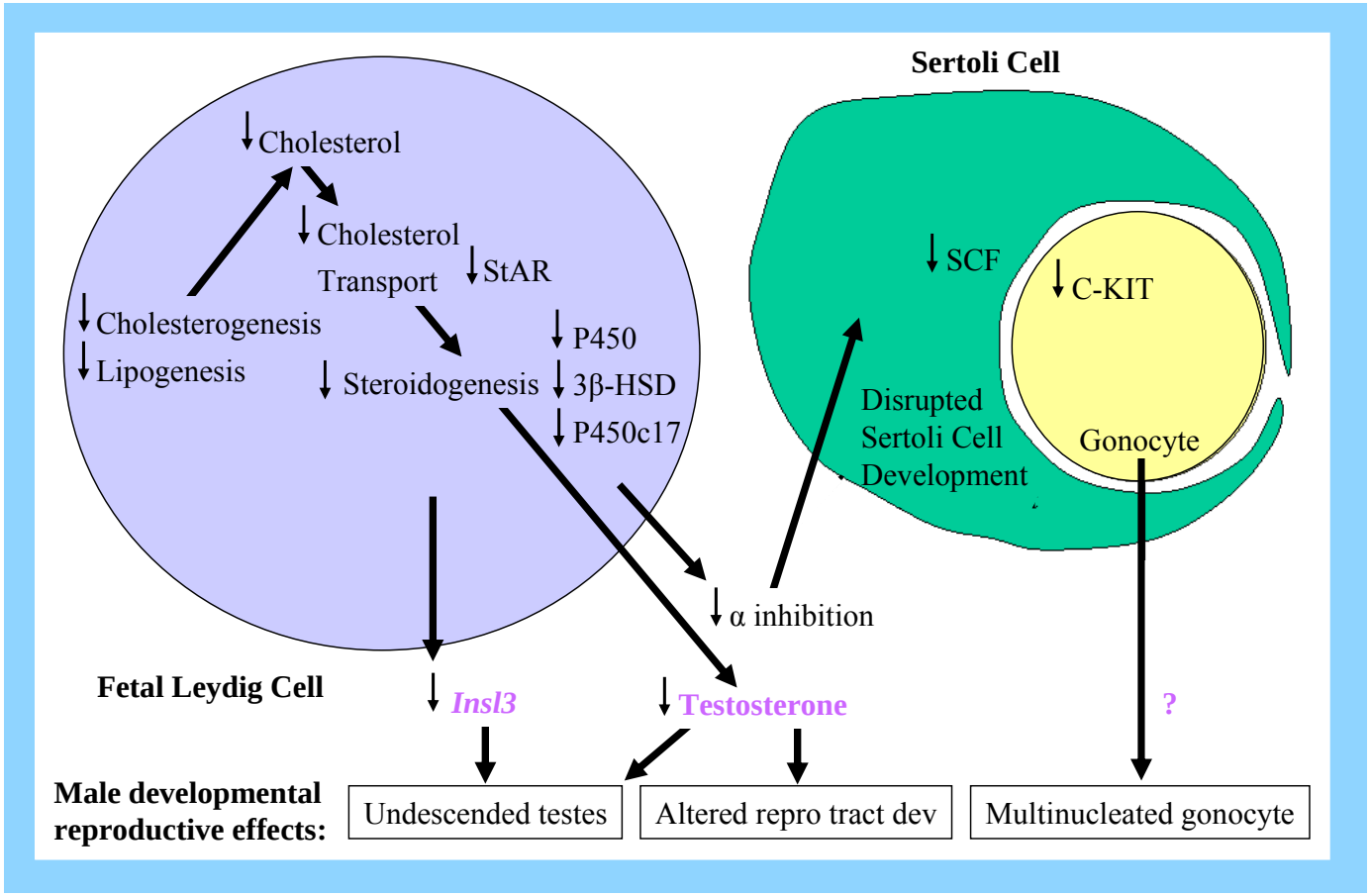

Figure 3-7. The proposed mechanism of action, defined as all steps between chemical exposure at the target tissue to expression of the outcome, for DBP. The steps shown are based on male reproductive developmental toxicity and toxicogenomic studies. Some of the affected pathways and individual genes whose expression was significantly affected by DBP exposure in multiple studies are included. By contrast, the proposed MOAs are shown in purple letters.

Source: Figure adapted from Liu et al. (2005), Thompson et al. (2004), Wilson et al. (2004), Barlow et al. (2003), and Shultz et al. (2001).

3.3. IDENTIFYING AND SELECTING QUESTIONS TO FOCUS THE DBP CASE STUDY

In reviewing the draft IRIS assessment and the DBP toxicogenomic data set, data gaps in the assessment were noted. We considered whether the DBP toxicogenomic data set could potentially address any of the gaps (see Figure 3-1). Four data gaps or questions of interest were identified.

Can the DBP toxicogenomic data set inform the

1) **Modes and mechanism of action for male reproductive developmental outcomes?** Not all of the male reproductive developmental outcomes after in utero DBP exposure are a consequence of reduced fetal testicular T (the critical effect selected in the current external review draft of the IRIS DBP assessment). For example, there is evidence that in utero exposure also reduces expression of *Insl3* mRNA. Additional MOAs may be identified by pathway analysis of the microarray data.

2) **Interspecies (rat to human) differences in MOA that could, in turn, inform the TD part of the UF_H?** There is evidence from toxicogenomic studies that a reduction in gene expression of some of the steroidogenesis genes underlies the observed reduction in fetal testicular T observed after in utero DBP exposure. Unfortunately, there are no genomic studies in appropriate human in vitro cell systems to make comparisons to in vivo rat MOA findings. Thus, the steroidogenesis pathway is one identified pathway affected by DBP exposure. Using available DNA sequence data and other methods, we would like to assess the rat-to-human conservation of the steroidogenesis pathway genes.

3) **Biologically significant level of reduction in fetal T?** The current external review draft of the IRIS DBP assessment selected a reduction in fetal testicular T as the critical effect. We considered whether the toxicogenomic data set could aid in determining the biologically meaningful level of T reduction.

4) **Dose-response assessment in risk assessment?** The microarray and RT-PCR studies have identified genes and pathways associated with the reduced fetal testicular T. Thus, there is the potential for evaluating these genes and pathways in a dose-response assessment.

Two questions (1 and 2 above) had the potential to be addressed utilizing the existing DBP toxicogenomics and other molecular data (i.e., for Question 2, "other molecular data" include DNA sequence data for comparison between rat and human steroidogenesis genes).

While of great interest, the available toxicogenomic data were not appropriate to address Questions 3 or 4 because of a lack of appropriate data. Questions 1 and 2 will be referred to in subsequent chapters as Case Study Question 1 and Case Study Question 2.

Subsequent steps include the evaluations of the toxicity data set for the male reproductive developmental effects after developmental exposure to DBP (Chapter 4) and the toxicogenomic data set (Chapter 5). Pathway analysis methods development was explored, and new analyses of some of the DBP microarray data were performed (Chapter 6) because analytical methods used for basic research studies may differ from analytical methods for application of these data to risk assessment. Chapter 4 follows with an in-depth evaluation of the DBP toxicity data set.

4. EVALUATION OF THE REPRODUCTIVE DEVELOPMENTAL TOXICITY DATA SET FOR DBP

This chapter presents the evaluation of the available toxicity data for the development of the male reproductive system following DBP exposure and the MOA(s) that contribute to the observed developmental outcomes of the male reproductive system. We used the compilation of the male reproductive toxicology literature cited in the draft U.S. EPA IRIS assessment (U.S. EPA, 2006a) as a starting point for our toxicology literature review for this case study. Each toxicology study was examined for the lowest dose and possible low-incidence effects in order to determine the full spectrum of male reproductive developmental effects. In a second evaluation, we used available information on MOA for each endpoint to identify "explained" and "unexplained" endpoints. The unexplained endpoints are one focus of the toxicogenomic data set evaluation, presented in Chapters 5 and 6.

An extensive toxicological data set exists for DBP that includes acute and subchronic studies in multiple species, multigeneration reproduction studies in rodents, and studies that assess developmental outcomes following in utero or perinatal/postnatal exposures. Following DBP exposure during the critical stages of development, the male reproductive system development is perturbed in rodent studies (Gray et al., 1999b, 2001; Mylchreest et al., 1998, 1999, 2000), and the MOA (see Chapter 2 and glossary for definition) of DBP for a number of these outcomes has been well established (David, 2006; Foster, 2005). The draft U.S. EPA IRIS assessment document (U.S. EPA, 2006a) utilized the alteration in fetal T levels, observed in Lehmann et al. (2004), as an endpoint for the derivation of acute, short-term, subchronic, and chronic reference values for DBP. This premise and conclusion were reviewed in the case study exercise, utilizing information from genomic studies that targeted and further elucidated the molecular events underlying these developmental outcomes (see Chapters 5 & 6). The intent of performing such an evaluation of the toxicology studies was to examine the possible usefulness of the toxicogenomic data in characterization of the MOA(s) that contribute to the adverse outcomes. We also examined the data for low dose or low incidence findings because such data may aid the interpretation of toxicological outcomes that can be misinterpreted as transient (e.g., AGD), or non-adverse due to low incidence or magnitude (e.g., statistically nonsignificant

incidences of gross pathology findings in male offspring reproductive organs, or alteration of fetal T levels).

4.1. CRITERIA AND RATIONALE FOR INCLUSION OF TOXICOLOGY STUDIES IN THE EVALUATION

Figure 4-1 illustrates the process of evaluating the toxicology data set for DBP, relevant to the goals of the case study. The first step in the process was the identification of studies that would be included for consideration in the case study. We identified a number of study selection criteria in Step 1. One criterion of prime importance was that the studies should include exposures to DBP during sensitive periods of male reproductive system development. Secondly, a no-observed-effect level (NOEL), lowest-observed-effect level (LOEL), or benchmark dose lower confidence limit (BMDL) would need to be identified for presumably adverse outcomes in the reproductive organs and/or function of male offspring. Additionally, the studies would need to be of adequate quality in order to establish confidence in the study conduct, methods, and results. These criteria, taken together, define a subset of the available toxicology studies that were considered possible candidates for determining the POD for derivation of reference values of various durations in the draft IRIS assessment document for DBP (see Tables 4-1, 4-2, and 4-3 in U.S. EPA, 2006a). These candidate study lists were considered during the External Peer Review of the IRIS document, conducted in July 2006, thereby providing a measure of confidence in their inclusiveness and veracity for the purpose of this case study. Though there are observable adverse effects on male reproductive system development in multiple species, the only available and relevant genomic studies with DBP (i.e., those that addressed effects on male reproductive system development following prenatal exposures) were conducted in rats. Table 4-1 lists the studies that were identified for inclusion as of July 2006. For each study, the following information was summarized: a description of the dose and exposure paradigm, the treatment-related outcomes observed at each dose level, and the experimentally derived NOEL and/or LOEL. The terms NOAEL and LOAEL are not used in this case study report, although these terms are commonly used in risk assessment. Some study reports do not specifically define NOELs or LOELs, and others do not address the issue of adversity of observed study outcomes. For that reason, Table 4-1 presents those outcomes that could be considered biomarkers of

effects on the male reproductive system that were reported by the study authors, without specific consideration or judgment of adversity.

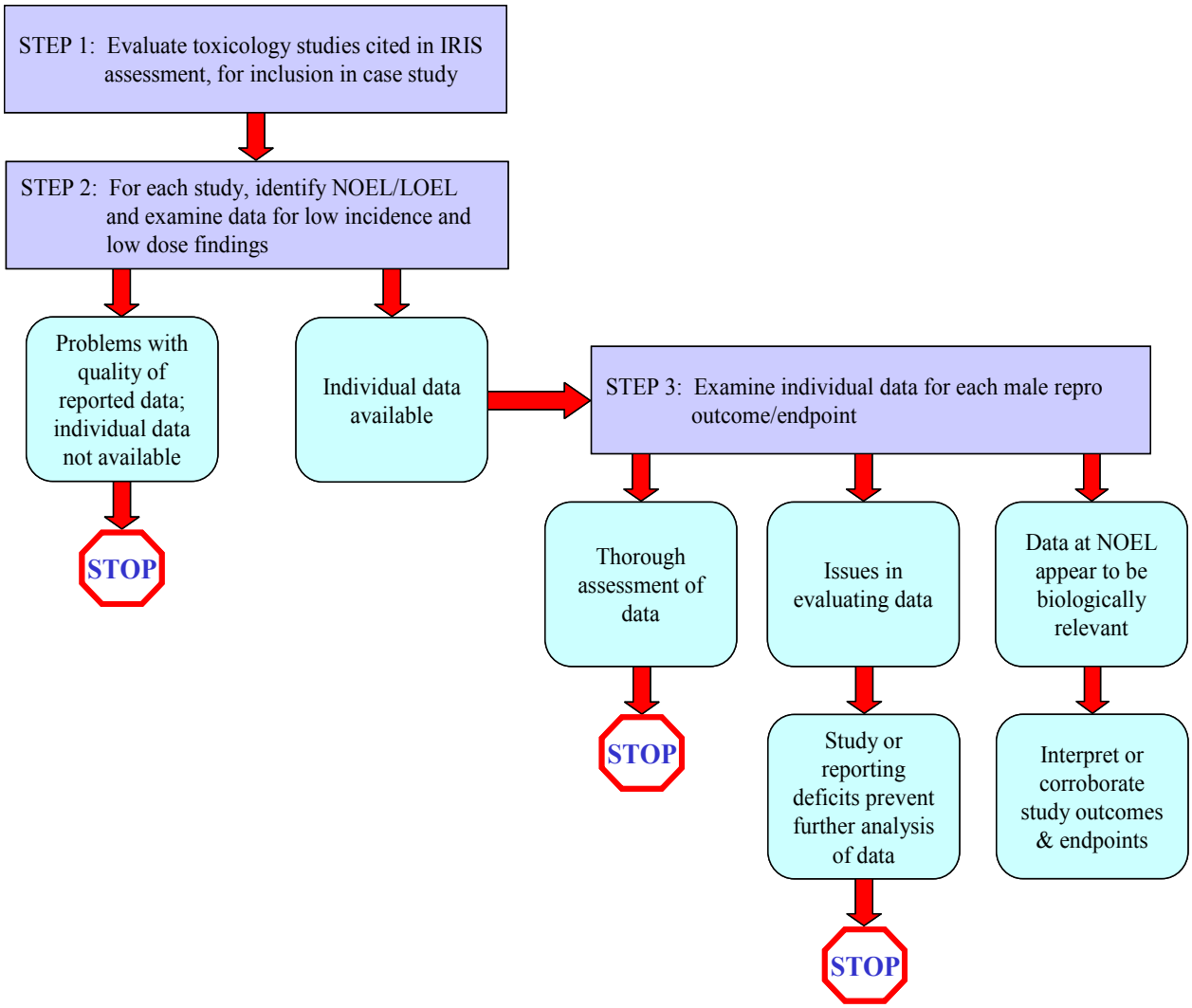

Figure 4-1. The process for evaluating the male reproductive developmental toxicity data set for low-dose and low-incidence findings.

Table 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

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	Repro NOEL mg/kg-d	Repro LOEL mg/kg-d
Barlow and Foster, 2003	Rat (SD); GD 12−21; 0 or 500 mg/kg-d	Large aggregates of Leydig cells, multinucleated gonocytes, & an increased number of gonocytes in fetal testes; a decreased number of spermatocytes on PND 16 & 21; epididymal lesions (decreased coiling of the epididymal duct, progressing to mild [PND 45], & then severe [PND 70] seminiferous epithelial degeneration).		500
Barlow et al., 2003	Rat (SD); GD 12−19; 500 mg/kg-d	Large aggregates of Leydig cells with lipid vacuoles.		500
Barlow et al., 2004	Rat (SD); GD 12−21; 0, 100, or 500 mg/kg-d	Testicular dysgenesis (proliferating Leydig cells & aberrant tubules); decreased AGD; areolae retention; small incidence of Leydig cell adenomas.	100	500
Bowman et al., 2005	Rat (SD); GD 12−19 or 21; 0 or 500 mg/kg-d	Marked underdevelopment of the Wolffian ducts (characterized by decreased coiling).		500
Carruthers and Foster, 2005	Rat (SD); GD 14−15, 15−16, 16−17, 17−18, 18−19, 19−20; 0 or 500 mg/kg-d	Decreased AGD; retained areolae & nipples; reduced epididymal weights, increased testes weight due to edema; malformations of the seminal vesicles, agenesis of various regions of the epididymis, small or flaccid testes; malformation of the coagulating gland.		500
Ema et al., 1998	Rat (Wistar); GD 11−21; 0, 331, 555, or 661 mg/kg-d	At 555 & 661 mg/kg-d, increased incidences of cryptorchidism & decreased AGD.	331	555
Ema et al., 2000b	Rat (Wistar); GD 15−17; 0, 500, 1,000, or 1,500 mg/kg-d	At 1,500 mg/kg-d, cryptorchidism observed in 80% of litters; at 500, 1,000, & 1,500 mg/kg-d, decreased AGD.		500
	Rat (Wistar); GD 12−14, or GD 20; 0, 1,000, or 1,500 mg/kg-d	At 1,500 mg/kg-d (GD 12−14), cryptorchidism observed in 50% of litters; at 1,000 & 1,500 mg/kg-d, decreased AGD.		1,000
Ferrara et al., 2006	Rat (Wistar); GD 13.5−21.5; 0 or 500 mg/kg-d	Delayed entry of gonocytes into quiescence, increase in gonocyte apoptosis, & subsequent early postnatal decrease in gonocyte numbers (exposures: GD 13.5−17.5); >10% increase in multinucleated gonocytes (exposures: GD 19.5−21.5).		500

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	Repro NOEL mg/kg-d	Repro LOEL mg/kg-d
Fisher et al., 2003	Rat (Wistar); GD 13−21; 0 or 500 mg/kg-d	Cryptorchidism, hypospadias, infertility, & testis abnormalities similar to human testicular dysgenesis syndrome; abnormal Sertoli cell-gonocyte interaction.		500
Gray et al., 1999b	Rat (Long-Evans) (P0); PND 21—adult; 0, 250, 500, or 1,000 mg/kg-d	At 250, 500, & 1,000 mg/kg-d, delayed puberty; at 500 & 1,000 mg/kg-d, reduced fertility related to testicular atrophy &reduced cauda epididymal sperm numbers.		250
	Rat (Long-Evans) (F1); GD 0−PND 21; 0, 250, 500, or 1,000 mg/kg-d	At 250 & 500 mg/kg-d, reproductive malformations (low incidences of hypospadias, testicular nondescent, & uterus unicornous); reduced fecundity.		250
	Rat (Long-Evans) (F1); GD 14 to PND 3; 0 or 500 mg/kg-d	Reduced AGD, retained nipples, permanently reduced androgen-dependent tissue weights.		500
Kim et al., 2004 **Ab**	Rat (SD); GD 10−19; 0, 250, 500, or 700 mg/kg-d	Decreased testes & accessory sex organ weight; delayed testis descent; increased expression of estrogen receptor in testes.		250 (presumed)
Kleymenova et al., 2004 **Ab**	Rat (strain not specified); GD 12−17, 19, 20; 0 or 500 mg/kg-d	Altered proliferation of Sertoli & peritubular cells; multinucleated gonocytes; changes in Sertoli cell-gonocyte interactions.		500[b]
Kleymenova et al., 2005a **Ab**	Rat (SD); GD 12−20; 0, 0.1, 1, 10, 30, 100, or 500 mg/kg-d	At 30 & 50 mg/kg-d, disruption of Sertoli-germ cell contact; at 50 mg/kg-d, Sertoli cell hypertrophy, decreased total cell number & number of seminiferous tubules; at 100 mg/kg-d, increased multinucleated gonocytes.	10	30
Kleymenova et al., 2005b	Rat (SD); GD 12−21; 0 or 500 mg/kg-d	Cytoplasmic changes in Sertoli cells with abnormal cell-cell contact with gonocytes, clustering of gonocytes in the middle of the tubules, altered morphometry of seminiferous tubules, clusters of interstitial cells, decreased number of tubular cross sections per testicular section; increased number of multinucleated gonocytes.		500

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	Repro NOEL mg/kg-d	Repro LOEL mg/kg-d
Lee et al., 2004	Rat (SD); GD 15 to PND 21; 0, 1.5, 14.4, 148, or 712 mg/kg-d (converted from 0, 20, 200, 2,000, and 10,000 ppm DBP in diet)	At 712 mg/kg-d, decreased percent males; decreased AGD & retained nipples, decreased relative testis weight; at 1.5, 14.4, 148, & 712 mg/kg-d, on PND 21, reduction in spermatocyte development, increased foci of aggregated Leydig cells, & decreased epididymal ductular cross section; at 148 & 712 mg/kg-d, at week 11, loss of germ cell development; at 1.5 mg/kg-d, degeneration & atrophy of mammary gland alveoli in males at 8−11 weeks of age.		1.5
Lehmann et al., 2004	Rat (SD); GD 12−19; 0, 0.1, 1, 10, 30, 50, 100, or 500 mg/kg-d	At ≥50 mg/kg-d, decreased fetal T concentration; at 500 mg/kg-d, a reduction in oil red O staining of lipids in fetal testes.	30	50
Liu et al., 2005	Rat (SD); GD 12−19; 0, 500 mg/kg-d	Significant reduction in AGD at GD 19.		500
Mahood et al., 2005	Rat (Wistar); GD 13.5−20.5; 0 or 500 mg/kg	Aggregation of fetal Leydig cells; reduced Leydig cell size; reduced T levels at GD 19.5 & 21.5 (early event in testicular dysgenesis); cryptorchidism; partial absence of epididymis at PND 90.		500
Mylchreest et al., 1998	Rat (SD); GD 30 to PND 20; 0, 250, 500, or 750 mg/kg-d	At 500 & 750 mg/kg-d, decreased AGD; at 250, 500, & 750 mg/kg-d, absent or underdeveloped epididymis, associated with testicular atrophy & germ cell loss, hypospadias, ectopic or absent testes; at 500 & 750 mg/kg-d, absent prostate & seminal vesicles, small testes, & seminal vesicles.		250
Mylchreest et al., 1999	Rat (SD); GD 12−21; 0, 100, 250, or 500 mg/kg-d	At 500 mg/kg-d, hypospadias; cryptorchidism; agenesis of the prostate, epididymis, & vas deferens; degeneration of the seminiferous epithelium; interstitial cell hyperplasia & adenoma; decreased weight of prostate, seminal vesicles, epididymis, & testes; at 250 & 500 mg/kg-d, retained areolae or thoracic nipples, decreased AGD; at 100 mg/kg-d, delayed preputial separation (attributed to highly affected litter, & not repeated in subsequent study).	100	250

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	Repro NOEL mg/kg-d	Repro LOEL mg/kg-d
Mylchreest et al., 2000	Rat (SD); GD 12−21; 0, 0.5, 5, 50, 100, or 500 mg/kg-d	At 500 mg/kg-d, decreased AGD, hypospadias, cryptorchidism, absent or partially developed epididymis, vas deferens, seminal vesicles, & ventral prostate; decreased weights of testes, epididymis, dorsolateral & ventral prostates, seminal vesicles, & levator anibulbocavernosus muscle; seminiferous tubule degeneration, focal Leydig cell hyperplasia, & Leydig cell adenoma; at 100 & 500 mg/kg-d, retained thoracic areolae or nipples in male pups.	50	100
Mylchreest et al., 2002	Rat (SD); GD 12−21; 0 or 500 mg/kg-d	In GD 18 & 21 fetuses, testicular atrophy, Leydig cell hyperplasia, enlarged seminiferous cords with multinucleated gonocytes; decreased testicular T; fewer epididymal ducts.		500
NTP, 1991	Rat (SD); continuous breeding (16 weeks) (gestation and lactation); 0, 80, 385, or 794 mg/kg-d in dams (converted from 0.1, 0.5, and 1.0 % DBP in feed)	F1 adults: At 80, 385, & 794 mg/kg-d: Increased incidence of absent, poorly developed, or atrophic testis & underdeveloped or absent epididymis. At 385 & 794 mg/kg-d: Increased incidence of seminiferous tubule degeneration. At 794 mg/k-d: Decreased mating, pregnancy, & fertility indices; decreased epididymal, prostate, seminal vesicle & testis weights; decreased cauda epididymal sperm concentration; decreased average spermatid count, total spermatid heads/testis or total spermatid heads /g testis; increased incidence of absent, small/underdeveloped/poorly developed, or atrophic penis, seminal vesicles, epididymis, & prostate; interstitial/Leydig cell hyperplasia; delayed testicular descent or cryptorchidism.		80
NTP, 1995 (some of this is also reported in Wine et al., 1997)	Rat (SD); continuous breeding (16 weeks) (gestation and lactation); 0, 80, 385, or 794 mg/kg-d in dams (converted from 0.1, 0.5, and 1.0% DBP in feed)	At 794 mg/k-d: Decreased mating, pregnancy, & fertility indices; decreased epididymal, prostate, seminal vesicle, & testis weights.	385	794

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	Repro NOEL mg/kg-d	Repro LOEL mg/kg-d
NTP, 1995	Rat (Fischer 344); perinatal and lactation plus 17 weeks; 0, 138, 279, 571, 1,262, or 2,495 mg/kg-d in dam? (converted from 0 or 10,000 ppm during gestation and lactation; 0, 1,250, 2,500, 5,000, 7500, 10,000, 20,000 ppm for 4 weeks PN; 0, 2,500, 5,000,10,000, 20,000, and 40,000 for last 13 weeks PN)	At 571, 1,262, & 2,495 mg/kg-d: Degeneration of germinal epithelium. At 1,262 & 2,495 mg/kg-d: Decreased testes & epididymal weights, fewer sperm heads per testis, & decreased epididymal sperm concentration.	279	571[c]
NTP, 1995	Rat (Fischer 344); perinatal and lactation plus 4 weeks; 0, 143, 284, 579, 879, or 1,115 mg/kg-d in dam (converted from 0, 1,250, 2,500, 5,000, 7,500, 10,000, and 20,000 ppm)	At 879 & 1,115 mg/kg-d: Moderate epididymal hypospermia in all males; at 579 mg/kg-d, mild epididymal hypospermia in 2 of 10 males.	284	579[d]
Plummer et al., 2005 **Ab**	Rat (strain not specified); gestation; 0 or 500 mg/kg-d	Decreased fetal T levels.		500
Shultz et al., 2001	Rat (SD), GD 12−21; 0 or 500 mg/kg-d	Decreased fetal testicular T & androstenedione; increased progesterone.		500
Thompson et al., 2004a	Rat (SD); GD 12−17, 12−18, or 12−19; 0 or 500 mg/kg-d	Decreased fetal T.		500
Thompson et al., 2005	Rat (SD); GD 19; 0 or 500 mg/kg-d	Decreased fetal T.		500
Wilson et al., 2004	Rat (SD); GD 14−18; 0 or 1,000 mg/kg-d	Decreased fetal T, expression of *Insl3*.		750

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	Repro NOEL mg/kg-d	Repro LOEL mg/kg-d
Zhang et al., 2004	Rat (SD); GD 1 to PND 21; 0, 50, 250, or 500 mg/kg-d	At 250 & 500 mg/kg-d, decreased AGD; underdeveloped epididymides; decreased epididymis or prostate weight at PND 70; decreased percent motile sperm & total sperm heads; degeneration of the seminiferous epithelium. At 500 mg/kg-d, cryptorchidism, absent epididymides, decreased total number of sperm.	50	250

Ab, Abstract only; AGD, anogenital distance; GD, gestation day; PND, postnatal day; Repro LOEL, lowest-observed-effect level for male reproductive system outcomes found in the study; Repro NOEL, no-observed-effect level for male reproductive system outcomes; T, testosterone. Note: These terms are used solely in a descriptive manner in this table, they may not reflect the terminology of the source study, and they are not intended to convey any regulatory implication.
[a]All studies used an oral route of exposure. Lee et al. (2004) and NTP (1995, 1991) exposed to DBP in the diet. All other studies used oral gavage.
[b]The abstract states that the effects were "dose dependent" but does not specifically indicate the LOEL.
[c]Overall, the study NOEL and LOEL are lower based on liver peroxisome activity.
[d]Overall, the study NOEL and LOEL are lower based on increased liver weight.

It is also noted that although BMDL values were calculated for specific developmental endpoints identified in Lehmann et al. (2004), Mylchreest et al. (2000), and the National Toxicology Program (NTP, 1995) (see draft IRIS document, Table 4-4), these values were not utilized as a POD for reference value derivation.

4.2. REVIEW OF THE TOXICOLOGY DATA SET

Figure 4-1 illustrates the stepwise approach taken in the evaluation of the toxicity studies, focusing on low-dose and low-incidence outcomes. First, for each toxicology study, we examined the data at the lowest dose levels (as defined by the study NOELs and LOELs) (Step 2). If there was any indication of insurmountable problems with the quality of the reported data (e.g., excessive variability, critical methodological concerns, lack of peer review as with abstracts, etc.), or if there were no individual animal data reported (as is often the case for poster abstracts as well as for many published studies, which only contain extracted summary data), the review of that study would be terminated. However, if individual data were available, the review could proceed (Step 3). The individual animal data were examined for evidence of reproductive system outcomes in the males. Although for most studies the exposures were only administered during the perinatal developmental period, we recognized that an adverse treatment-related outcome might be identified at any life stage that was assessed in the study. There were three possible courses that the data review could take from this point forward. In cases where problems were identified in the data, we attempted to analyze the extent of the issues and determine the ability to move forward with the study analysis. In some cases the analysis stopped at this point, due to deficits in the study data or to inadequate reporting of individual animal data. However, if the data in the report appeared to be thoroughly assessed, then the study outcomes and endpoints were examined. Alternatively, in some cases where adequate individual study data were available for analysis, further examination of the study could identify effects at the lowest dose levels that had been considered biologically irrelevant in the original review, but it might require further consideration. At any point in this stepwise process that data were deemed insufficient to proceed further, we identified research needs (discussed in Chapter 7).

To begin the characterization and evaluation of the published studies according to this stepwise model, important aspects of each study protocol, conduct, and reporting were first

summarized (Table 4-2). Examination of this table demonstrates that approximately half the studies that were selected for analysis (i.e., 14 of 29) were limited to a single-dose group, which eliminated them from further examination for lower-dose level effects. It is also important to note that individual animal data were reported in only 2 of the 29 studies, thereby severely limiting, and in some cases even preventing, more rigorous evaluation of the study findings. These two characteristics alone tend to overshadow any of the other listed study attributes that might contribute to confidence in study findings (i.e., evidence that the study was conducted according to quality laboratory standards, description of statistical analysis of the data, and/or specific information regarding the number of litters and offspring assessed, which would provide an indicator of statistical power). Of the studies listed, only the study conducted by the NTP (1995, 1991) was considered suitable for extended examination.

In order to create a profile of outcomes to the male reproductive system following developmental exposures, which might then serve as a baseline for further comparison and analysis of toxicological findings across the studies, a list of observed effects was compiled (Table 4-3). The content of this list is very clearly defined by the study protocols, both in terms of what endpoints were examined in each study and when (i.e., at what life stage) they were examined. For some endpoints, the precise GD or postnatal day (PND) of evaluation may even be critical. For example, fetal T should peak at approximately gestation day (GD) 18, so assessments made at earlier or later time points may be less sensitive in detecting adverse outcomes, and the effects will not be directly comparable across fetal ages. Decreases in T levels may not be observed postnatally unless treatment is continued or if testicular malformations disrupt T level (which is a different mechanism of perturbation than alterations to the steroidogenic pathway). In neonates, examination for nipple retention is generally conducted at around PND 13, when the structure is readily visible but before it is obscured by hair growth.

Cryptorchidism, even though present at birth, may not be readily observable in neonates until they reach the age of PND 16−21 (and of course, it should be detectable at postweaning ages and in adults). Preputial separation (PPS) delays can only be observed at the time of sexual maturation, which, in the male Sprague-Dawley rat, occurs at approximately PND 42; therefore, this effect cannot be detected at an earlier life stage, nor will it be observed in sexually mature adults. On the other hand, sperm alterations (count, morphology, or motility) and perturbations

in male fertility can only be assessed in adult males, not in immature individuals at earlier life stages.

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

Study	>One high dose	Individual data publicly available	Stat analysis method reported	Study conduct level reported	Number evaluated/group	
					Litters	Offspring
Barlow and Foster, 2003			✓	✓	1−9[a]	7−60[a]
Barlow et al., 2003		✓ subset[b]	✓	✓	NR	3
Barlow et al., 2004	✓		✓	✓	8−11[a]	35−74[a,c]
Bowman et al., 2005			✓	✓	18	All male fetuses
Carruthers and Foster, 2005			✓	✓	1−14[d,e]	1−91[e]
Ema et al., 1998	✓		✓		11 DBP treated	AGD: NR; crypt.: 144
Ema et al., 2000b	✓		✓		73 DBP treated	~770[f]
Ferrara et al., 2006			✓	✓	"in most instances" ~3−6	1−3/litter[g]
Fisher et al., 2003			✓	✓	NR	Testis wt: 5−10 animals/age group (4); hyp. & crypt.: 10 adults
Gray et al., 1999b	✓ PPS only		✓		4 (LE); 8 (SD)	LE: 30 male pups; 13 adult males SD: 48 male pups; 17 adult males[h]
Kim et al., 2004 **Ab**	✓				NR	NR
Kleymenova et al., 2004 **Ab**	✓				NR	NR
Kleymenova et al., 2005a **Ab**	✓				NR	NR
Kleymenova et al., 2005b			✓	✓	3	14−21 pups/evaluation
Lee et al., 2004	✓		✓		6−8	11−20 adults
Lehmann et al., 2004	✓		✓	✓	1−4	3−4 fetuses/group
Liu et al., 2005			✓	✓	3	3 fetuses/litter

Table 4-2. (continued)

Study	>One high dose	Individual data publicly available	Stat analysis method reported	Study conduct level reported	Number evaluated/group	
					Litters	Offspring
Mahood et al., 2005			✓	✓	2−7	NR
Mylchreest et al., 1998	✓		✓	✓	7−10	All males/litter
Mylchreest et al., 1999	✓		✓	✓	10	All males/litter
Mylchreest et al., 2000	✓		✓	✓	11−20	All males/litter
Mylchreest et al., 2002			✓	✓	5−6	23−49 fetuses
NTP, 1995, 1991	✓	✓	✓	✓	20	All pups/litter in-life thru necropsy; histopath: 10/selected groups
Plummer et al., 2005 **Ab**					NR	NR
Shultz et al., 2001			✓	✓	3	1 male/litter
Thompson et al., 2004a			✓	✓	4	1 male/litter
Thompson et al., 2005			✓	✓	4	3 fetuses/litter
Wilson et al., 2004			✓	✓	3	All males/litter
Zhang et al., 2004	✓		✓		14−16	20 pups/group

Ab, Abstract only; LE, Long Evans; NR, Not reported; PPS, preputial separation; Y, present.
[a]Litters and pup numbers not reported for AGD and areolae retention.
[b]Data for three individual animals were reported for LC and Sertoli cell staining. The other results are not reported in this table because they were from toxicogenomic studies (see Chapter 5).
[c]57−100% of these pups survived to necropsy so for malformations that required necropsy, the number of pups is less than shown.
[d]Reported mean litter size for Table 1.
[e]Litters for AGD were the statistical unit; neither litter nor pup numbers for AGD were reported.
[f]Number derived from the mean number of live fetuses/litter.
[g]In some cases, data from two experiments were combined.

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

Findings	Life stage of animals (rats) at observation		
	Fetus	Neonate through puberty	Adult
Decreased T	✓	✓	✓
Malformations	✓	✓	✓
Decreased AGD		✓	✓
Hypospadias		✓	✓
Retained nipples/areolae		✓	✓
Cryptorchidism		✓	✓
Delayed PPS		✓	
Organ weights		✓	✓
Histopathology of male reproductive organs	✓	✓	✓
Abnormal sperm			✓
Decreased fertility			✓

T, Testosterone; AGD, Anogenital distance; PPS, Preputial separation

Using the list in Table 4-3 as a guide, a more extended analysis was conducted for each of the selected studies. Table 4-4 presents the detailed results. In this table, the various observed outcomes are arrayed across three general life stage categories: prenatal (i.e., observations conducted in fetuses), neonatal through puberty (i.e., observations conducted in pups), and adult (i.e., observations conducted in young, sexually mature animals). These life stage categories do not represent the period of exposure for the study. While all studies include exposures during late gestation (i.e., during the critical window of male reproductive system development), some studies also maintained exposures during later life stages. For reference, Table 4-1 provides general descriptions of exposure durations.

Table 4-4 summarizes the outcomes and presents a broad representation of positive and negative observations in a manner that demonstrates that not all relevant endpoints were evaluated at all life stages or even in each study. To facilitate summarization of the myriad

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

	Fetus		Neonate through puberty								Adult									
	↓ T[a]	Histo-path[b]	↓ AGD[c]	Hyp[d]	Ret. nip/ areolae[e]	Crypt[f]	Del. PPS[g]	↓ Org wt[h]	Histo-path[b]	↓ T[a]	Malf[i]	↓ Org wt[h]	Histo-path[b]	Ab. Sperm[j]	↓ Fert[k]	Hyp[d]	Ret. nip/ areolae[e]	Crypt[f]	Δ AGD	↓ T[a]
Barlow and Foster, 2003		✓	✓	✓	✓	✓	✓—		✓		✓	✓[l]	✓	✓		✓		✓		
Barlow et al., 2003		✓																		
Barlow et al., 2004			✓		✓						✓	✓[m]	✓			✓[n]	✓	✓	✓↓	
Bowman et al., 2005		✓[o]																		
Carruthers and Foster, 2005			✓[p]		✓[q]						✓	✓	✓			—	✓[r]	—	✓[s]	
Ema et al., 1998		✓	✓[t]			✓[t]					?									
Ema et al., 2000b		?	✓[t]			✓[t]														
Fisher et al., 2003	✓	✓						✓		✓ /—[u]	✓	✓	✓	✓	✓	✓		✓		—[u]
Gray et al., 1999b			✓		✓		✓[v]				✓	✓	✓	✓[w]	P0 males	✓	✓	✓		P0 males
Kim et al., 2004 **Ab**				✓		✓		✓	—	—[x]	—	—	—							✓[x]
Kleymenova et al., 2004 **Ab**		✓																		
Kleymenova et al., 2005a **Ab**		✓																		
Kleymenova et al., 2005b		✓		—		—			✓											
Lee et al., 2004			✓	—[y]	✓	—[y]	—	—	✓		—[y]	✓	✓							
Lehmann et al., 2004	✓																			
Liu et al., 2005			✓[z]																	
Mahood et al., 2005	✓	✓				✓[aa]		NR	✓		✓	NR	✓			NR		✓		
Mylchreest et al., 1998			✓	✓		✓	—				✓	✓		—[bb]		✓		✓		
Mylchreest et al., 1999			✓	✓	✓	✓	✓				✓	✓	✓			✓		✓		
Mylchreest et al., 2000			✓	✓	✓	✓	—				✓	✓	✓			✓		NR		
Mylchreest et al., 2002	✓	✓	NR[z]																	

Table 4-4. (continued)

	Fetus		Neonate through puberty								Adult									
	↓ T[a]	Histo-path[b]	↓ AGD[c]	Hyp[d]	Ret. nip/ areolae[e]	Crypt[f]	Del. PPS[g]	↓ Org wt[h]	Histo-path[b]	↓ T[a]	Malf[i]	↓ Org wt[h]	Histo-path[b]	Ab. Sperm[j]	↓ Fert[k]	Hyp[d]	Ret. nip/ areolae[e]	Crypt[f]	Δ AGD	↓ T[a]
NTP, 1991				✓							✓	—?	✓	✓	✓	✓		✓		
Plummer et al., 2005	✓					✓[cc]														
Shultz et al., 2001	✓		NR[z]																	
Thompson et al., 2004a	✓																			
Thompson et al., 2005	✓																			
Wilson et al., 2004	✓																			
Zhang et al., 2004			✓	—		✓		—			✓	✓	✓	✓[bb]		—		✓		

Y, Observed; —, Not observed; white box, Not evaluated; shaded box, Evaluated; NR, Not reported, although the study indicates that the endpoint was evaluated. **Ab**, Abstract only; PPS, preputial separation.

[a]Decreased testicular testosterone (T) should peak at PND 18; Fisher et al. (2003) also assessed plasma T levels postnatally and in adults, but the relevance of their findings is unclear.

[b]Histological changes—Leydig cell hyperplasia (aggregation); multinucleated gonocytes; Wolffian duct increased coiling (can be measured in fetus, neonate through puberty, or adult).

[c]Decreased AGD; or Δ for change in AGD.

[d]Hypospadias.

[e]Retention of nipples.

[f]Cryptorchidism (can observe between PND 16−21 and older).

[g]Delayed preputial separation (normally observed ~PND 42).

[h]Organ weight decreases (see list below); a decrease in organ weight in at least one reproductive organ was observed.

[i]Malformations—ventral/dorsal/lateral prostate, seminal vesicles, androgen dependent muscles, (accessory sex organs) epididymis, vas deferens external genitalia, cryptorchidism, small or flaccid testes.

[j]Sperm changes—count, motility, morphology.

[k]Decreased fertility.

[l]Enlargement of the seminiferous cords was observed at PND 19−21.

[m]In addition to the observed decreases and absences of male reproductive organs, "occasional enlargement" of the testes was observed only in the 500 mg/kg-d group.

[n]Assessed in adult animals at PND 180, 370, and 540. Hypospadias only observed in the 500-mg/kg-d group.

[o]Wolffian ducts smaller, more fragile, adipose tissue surrounding duct was more gelatinous, and decreased coiling.

[p]Assessed at PND 1 and 13. Reduction in AGD observed in animals exposed to DBP on GDs 16 & 17, GDs 17 & 18, or GD 19s & 20; no change in AGD in animals exposed GD 14 and 15.

Table 4-4. (continued).

[q]Assessed on PND13; assessed on an individual animal basis, significant increase in nipple retention was observed after dosing on GD 15−16; 16−17; 17−18; or 19−20.
[r]Assessed at PND 90; significant increase in nipple retention only for males dosed GD 16−17 (individual animal basis).
[s]Increased AGD seen in animals exposed GD 16 and 17; no observable change in animals exposed GDs 17 & 18, GDs 19 & 20, or exposed GDs 14 & 15.
[t]AGD and cryptorchidism were assessed in fetuses on GD21. Exposed pregnant dams were sacrificed on day 21, and live fetuses were removed.
[u]Assessed blood plasma T levels significantly reduced on PND 25 but not on PND 4, 10, or in adult.
[v]Delayed PPS only reported for parental generation (P0) males exposed from weaning through to puberty.
[w]Reduced epididymal sperm numbers; not necessarily abnormal sperm.
[x]Evaluated T levels at 31 and 42 days (not fetus) and found decreased at 42 days.
[y]It is presumed that specific malformations would have been observed if present based on the study design and methods.
[z]Examined in GD 19 or 21 fetuses.
[aa]Observed at PND 25 and 90; nonscrotal testes were not evaluated histopathologically.
[bb]Only motility was evaluated in Mylchreest et al. (1998); in Zhang et al. (2004), sperm number, motility, and morphology were evaluated, but only count was affected.
[cc]Study mentions that adult cryptorchidism was observed, but study methods do not indicate that offspring were retained until adult age.

individual study findings, information was often combined by category (e.g., "histopathology" includes a broad variety of outcomes in various reproductive organs), and for the sake of brevity, the minute details and nuances of the study design and observations, although quite interesting, are not typically presented. In a few cases, negative outcomes presented in the table are extrapolations based upon the presumption that specific findings would have been observed if they were present. For example, with methods that include detailed external and internal (macropathology) examination of pups and/or adults, the absence of reported malformations at either of these life stages was presumed to indicate that no gross malformations were observed because they should have been readily detectable (e.g., Lee et al., 2004).

Tables 4-1, 4-3, and 4-4 clearly illustrate that the study protocols varied quite extensively. In general, with the exception of the NTP studies, the protocols were not designed to conform to a particular regulatory guideline. Rather, the majority of the studies were focused research efforts that were verifying and/or expanding upon previously observed outcomes; therefore, the differences across study methods are understandable. As a result, the apparent lack of consistency in male reproductive system observations across studies is generally attributable to differences in protocol design and implementation. Some examples are discussed in detail as follows:

- Although these studies all utilized exposures during late gestation (i.e., a critical period of male reproductive system development in the rat), the specific endpoints that were assessed and/or the life stages at which endpoints were examined varied extensively across the studies. Obviously, treatment-related alterations of life-stage-specific events require examination during the most appropriate or optimal life stage (for example, increased multinucleated gonocytes can only be observed in fetal testes, delays in PPS can only be observed in juvenile animals at the time of sexual maturation, and disturbances in reproductive function can only be observed in sexually mature adults). Other permanent structural abnormalities may be detected across multiple life stages (e.g., hypospadias or cryptorchidism could theoretically be observed in late gestation fetuses, in adolescents, and in adults). For some outcomes, it is difficult to predict the optimal time point for evaluation. For example, DBP-related decreases in the ER were observed at 31 days but not at 42 days (Kim et al., 2004).

- It is important to realize that not all available offspring are evaluated in every study; therefore, identification of adverse outcomes may rely in part on sampling protocols and the statistical power of the sample size for detection of rare or low-incidence events. Calculations of statistical power are rarely provided in study reports.

- In some cases, apparent differences in studies may result because the report contains an insufficient level of detail on a particular endpoint or life stage—often because the emphasis of the scientific review lies in a slightly different direction. For example, if high doses of DBP are administered during sensitive periods of male reproductive system development, and the males are maintained on study and terminated as adults, at which time histopathological evaluation is performed, it might be assumed that various male reproductive system malformations and/or cryptorchidism would have been present in some of the males at necropsy. Yet, these findings may not be reported because the histopathological findings are the primary focus of the investigation and/or the publication (e.g., Lee et al., 2004).

- In other situations, the description of the findings at various life stages may vary. For example, evidence of cryptorchidism may be described as "testis located high in the abdomen" in a fetus, as "undescended testis(es)" in an adolescent rat, or as "unilateral testis" upon noninvasive clinical examination of an adult. To some extent, this lack of consistency in terminology may result from laboratory Standard Operating Procedures that direct technical staff to avoid the use of diagnostic terminology.

Overall, in spite of numerous differences in the study designs, the toxicological profile for DBP clearly demonstrates that exposure to DBP during critical stages of male reproductive system development can result in adverse structural and functional reproductive outcomes. When specific critical aspects of study design and implementation were similar, consistent outcomes were almost universally observed. The WOE embodied by the data described above is further supported by studies in rats that demonstrated similar incidences of cryptorchidism and decreased AGD in male pups of dams treated with either DBP or MBP, the metabolite of DBP (Ema and Miyawaki, 2001). The ability of MBP to cross the placenta and reach the fetus has also been conclusively demonstrated (Fennell, 2004; Saillenfait et al., 1998), and these two TK events (metabolism and placental transport) are key to the MOA of reduced fetal testicular T (David, 2006). Available toxicogenomic data, described elsewhere in this case study document, further elucidate the MOA(s) of DBP in producing adverse effects on male reproductive system development and are an important consideration in the WOE analysis of the toxicological profile.

In the selected DBP toxicology study data set, the presentation of extensive individual offspring data was limited to the NTP (1991) study conducted as a reproductive assessment by continuous breeding (RACB) in SD rats. The individual data from this study were carefully examined in order to confirm the NOEL and LOEL described in the study report. This analysis was conducted under the presumption that statistical and/or biological significance noted in the

summary compilations of male reproductive system outcomes might not identify low incidence effects in individual offspring at lower dose levels. To further aid the identification of treatment-related outcomes, the male reproductive system outcomes were grouped by organ instead of individual animal. This analysis revealed apparently treatment-related findings in the testis and epididymis of F1 male offspring, as summarized in Table 4-5. At the highest dose tested (794 mg/kg-d, equivalent to 1.0% DBP in the diet), additional findings in the male reproductive organs of F1 offspring included single incidences of (1) underdeveloped prepuce; (2) mild secretion and severe vesiculitis of the prostate; (3) a mass on the testis; and (4) a focal granuloma with fluid and cellular degeneration in the epididymis; these findings were not observed at the lower dose levels. Understandably, the findings at the low- and mid-dose groups were not originally interpreted as being treatment related (Wine et al., 1997; NTP, 1991). However, consideration of MOA information for DBP, including toxicogenomic data, has resulted in a more conservative interpretation of the data both by NTP researchers (Paul Foster, personal communication, 2008) and by the U.S. EPA IRIS program (U.S. EPA, 2006a). Consequently, further analysis of individual offspring data in the current case study did not identify any additional sensitive toxicological outcomes; the study LOEL was confirmed to be the lowest treatment level tested in the NTP RACB study (80 mg/kg-d).

4.3. UNEXPLAINED MODES OF ACTION (MOAS) FOR DBP MALE REPRODUCTIVE TOXICITY OUTCOMES

Figure 3-6 illustrates the broad conceptual approach for consideration and interpretation of toxicogenomic and toxicology data to inform MOA. The toxicogenomic data can be evaluated to identify altered genes, gene products, and pathways; this information can lead to a more complete understanding of the mechanism of action or MOA(s) for the chemical toxicity. From the opposite perspective, the toxicity data can provide information

Table 4-5. Incidence of gross pathology in F1 male reproductive organs in one continuous breeding study with DBP[a]

Gross finding[b]	Dose (% in Diet)			
	0	0.1	0.5	1.0
Testis: absent, poorly developed, atrophic, undescended	0/20	1/20	1/20	6/20
Penis: small/underdeveloped	0/20	0/20	0/20	4/20
Epididymis: underdeveloped/absent	0/20	1/20	1/20	12/20

[a]Incidences were compiled from reported individual animal macroscopic pathology data; statistical analysis was not performed.
[b]Some animals have more than one type of malformation, and these animals were counted separately for each of the three outcome categories.

Source: (NTP, 1991).

critical to identifying the relevant MOA(s) involved in the toxicological outcomes, and thereby inform the interpretation of gene alterations and relevant pathways.

Consideration of the MOA for each outcome, in conjunction with pathways identified in the toxicogenomic data set, may either help to corroborate known or hypothesized MOAs or suggest the existence of other potential MOAs (see Figure 4-2). For the DBP case study, Table 4-6 presents a compendium of the specific findings noted in the male reproductive system following exposures at critical windows of development. Each outcome is associated with specific known MOAs. While reduced fetal testicular T and reduced *Insl3* signaling can be linked to some of the observed outcomes on the basis of available data, potential key events cannot be specifically identified for other outcomes.

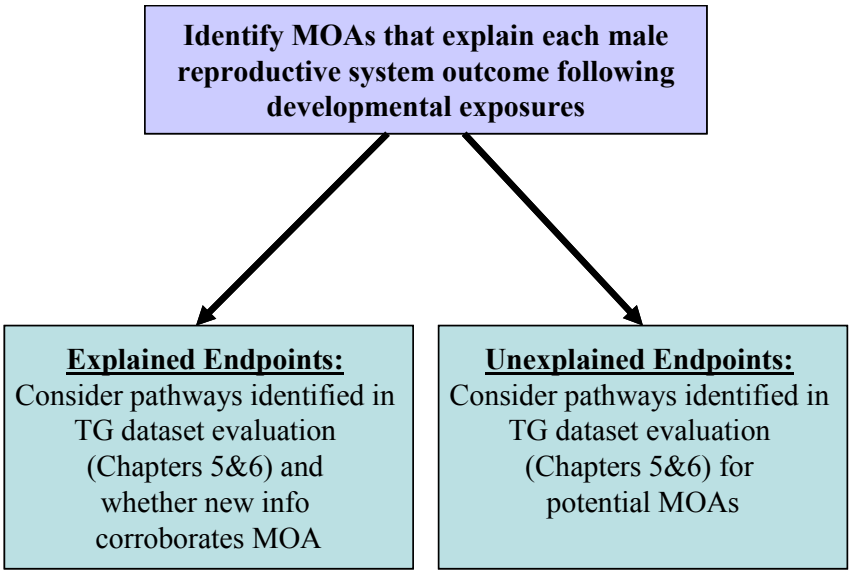

Figure 4-2. The process for evaluating the MOA for individual male reproductive developmental outcomes.

Table 4-6. Effects in the male reproductive system after in utero DBP exposure, and MOAs[a] that explain the affected endpoints

Organ/ Function	Effect	MOA	
		Reduced fetal testicular T	Reduced *Insl3* signaling
Testes	Multinucleated gonocytes; increased number of gonocytes in fetal testes	?[b]	?[c]
	Altered proliferation of Sertoli and peritubular cells; fewer Sertoli cells	?[b]	?[c]
	Gonocyte apoptosis increase; early postnatal decrease in gonocyte number	?[b]	?[c]
	Abnormal Sertoli cell-gonocyte interaction	?[b]	?[c]
	Small incidence of Leydig cell adenomas, aggregates, and hyperplasia	✓	?[c]
	Decreased number of spermatocytes or cauda epididymal sperm concentration.	✓	✓[d]
	Small or flaccid; other abnormalities; decreased weight	✓	✓
	Increased weight due to edema	?[e]	?
	Decreased number or degeneration of seminiferous cords/tubules; altered morphology; degeneration of the epithelium; enlarged cords/tubules	?[b]	?[c]
	Testes descent: none (cryptorchid) or delayed	✓[f]	✓[f]
Gubernacular ligament	Gubernacular ligament development effects: agenesis or elongation	X	✓
Epididymis	Lesions and agenesis; partial to complete absence; decreased epididymal ductular cross section	✓	X
	Reduced weights	✓	✓

Table 4-6. (continued)

Organ/ Function	Effect	MOA	
		Reduced fetal testicular T	Reduced *Insl3* signaling
Mammary gland	Nipple and/or areolae retention in males	✓	**X**
	Degeneration and atrophy of alveoli in males	**?**[b]	**X**
Wolffian ducts	Underdeveloped	✓	**X**
Seminal vesicles	Malformations or absent; decreased weight	✓	**X**
Coagulating gland	Malformations	✓	**X**
Penis	Small, underdeveloped	✓	**X**
	Hypospadias	✓	**X**
	Delayed preputial separation	✓	**X**
Accessory sex organ	Decreased weight	✓	**X**
Prostate	Decreased wt or absent	✓	**X**
Vas deferens	Agenesis	✓	**X**
Levator anibulbocavernosus muscle	Decreased weight	✓	**?**[c]
Male/female ratio	Decreased % male offspring as determined by AGD at birth	✓	**X**
Perineum	Decreased AGD	✓	**X**
Repro function	Infertility	✓	**Y**[d]

AGD, anogenital distance; ?, Current data indicate that it is unlikely the MOA; **Y**, Current weight of evidence of the data support this MOA leading to the effect**; X**, Current weight of evidence of the data indicate that this MOA is not the MOA for this outcome.

[a]MOA is defined as one or a sequence of key events that the outcome is dependent upon (see glossary).

[b]Reduced fetal testicular T may play a role, but current data indicate that reduced T is not solely responsible for this outcome.

[c]The *Insl3* knockout mouse phenotype suggests that *Insl3* is specifically required for gubernacular ligament development and, therefore, testis descent in mice since these mice do not have other defects.

Table 4-6. (continued)

[d]Decreased fertility in males is a result of reduced *Insl3* signaling since reduced *Insl3* signaling leads to undescended testes, which, in turn, reduces sperm count (presumably by increasing the temperature) and can cause infertility.
[e]In some animals, increased weight, due to edema, can result in animals that have epididymal agenesis, which is a consequence of reduced testosterone (T).
[f]*Insl3* signaling is required for development of the gubernacular ligament and through this mechanism—the 1st stage of testis descent from the kidney region to the inguinal region. Testosterone is required for the 2nd stage of testis descent, from the inguinal region to the scrotum (reviewed in Klonisch et al., 2004). After in utero DBP exposure, the cryptorchid phenotype resembles the *Insl3* knockout. A delay in testis descent can result from reduced *Insl3* and T.

4.4. CONCLUSIONS ABOUT THE TOXICITY DATA SET EVALUATION: DECISIONS AND RATIONALE

The review of the toxicology data set identified a number of issues and limitations that are evident in the study descriptions and endpoint summaries presented in this chapter. These include the following:

- *Lack of dose-response information:* A number of studies conducted with DBP used a single high-dose treatment level (often at 500 mg/kg-d) in order to produce readily observable adverse outcomes to male reproductive system development that could be examined. In such studies, the absence of lower-dose levels prevents the evaluation of dose-dependent responses and does not allow the identification of study-specific NOELs or LOELs. While this approach is useful for hazard characterization, it does not facilitate other aspects of risk assessment (e.g., dose-response assessment or risk characterization). Thus, studies utilizing a single high-dose level may provide important information for a WOE assessment of the toxicology profile, but they have diminished usefulness in identifying outcomes for use in risk calculations at environmentally relevant doses.

- *Insufficient information on study methods:* Even though every study report includes a section on study methods, there can be a great deal of unevenness in the amount of detailed information provided. Consequently, important questions may arise during study review that cannot be readily resolved. In some cases, this can have an impact on individual study interpretation or on conclusions that rely upon a thorough WOE evaluation of the data set.

- *Unavailable individual outcome data*: A full range of individual animal data is seldom included in studies published in the open literature and is almost never available when the only available publication is a presentation abstract. Conversely, individual animal data are generally included in toxicology reports generated in response to a regulatory mandate or conducted by a federal agency (e.g., NTP). The availability of individual animal data can be quite important in interpreting the study findings, because it can reveal problems or inadequacies in the data, but it can also help identify low incidence

adverse outcomes. In the case of DBP, the individual offspring data presented in the NTP study report (1991) include alterations in the reproductive system of F1 males that had been exposed during development. These findings are similar to outcomes identified at higher-dose levels, are consistent with the proposed MOA, and, consequently, are used to establish a LOEL for the study.

- *Protocol limitations:* Unless studies are designed to meet the recommendations of a standardized testing protocol (e.g., NTP or U.S. EPA/Office of Prevention, Pesticides and Toxic Substances reproductive toxicity study guidelines), there may be a high degree of variability among the protocols used for testing any one chemical. Between two studies, there can be differences in the treatment regimen or in the assessment of outcomes that render them incomparable. DBP provides a good example of a chemical that targets a very specific critical prenatal window of reproductive system development in males, and results in adverse outcomes that could go unidentified if the appropriate endpoint(s) are not assessed at the optimal life stage or time point.

- *Specific study's limitations:* Even when a study design optimizes the detection of adverse outcomes from chemical treatment, there may be challenges in study analysis and interpretation. Such is the case with the NTP study (1995, 1991) on DBP, which was conducted in several phases and reported both in the open literature (Wine et al., 1997) and by the Institute that conducted the experiments.

The analysis of the toxicology data in this chapter has provided a firm basis for expanded consideration of the toxicogenomic data for DBP as depicted in Figure 3-6. The extensive analysis of the toxicology data set and consideration of MOA(s) provide a source of information for use in phenotypic linking of known and potential MOAs. The available toxicogenomic studies for DBP are evaluated in Chapter 5. The genes and pathways underlying the endpoints with well established or unexplained MOAs are utilized in Chapter 5, where consistency of findings for altered genes and pathways are evaluated, and, in Chapter 6, where the new pathway analyses are presented.

5. EVALUATION OF THE DBP TOXICOGENOMIC DATA SET FROM THE PUBLISHED LITERATURE

This chapter presents an evaluation of the DBP toxicogenomic data set from the published literature. The toxicogenomic studies include nine published RT-PCR and microarray studies in the rat after in utero DBP exposure. We evaluated the toxicogenomic data set for (1) the consistency of findings from the published studies, and (2) whether additional pathways affected by DBP in utero exposure could possibly explain the testis endpoints for which there is not an established MOA (these "unexplained" endpoints were identified in Chapter 4). The DBP genomics data set includes nine papers published through July 2008. The microarray studies all reported DBP doses of 500 to 1000 mg/kg-d during the critical window for male reproductive development, which is during late gestation and correlates with the time period of maximal T production. The chapter first discusses the methodologies utilized in the nine studies and provides a brief overview of each study. The chapter then presents an evaluation of the consistency of the findings or WOE for the microarray, RT-PCR, and protein studies performed in the rat testes. The findings of one DBP dose-response RT-PCR study of Lehmann et al. (2004) are discussed. The chapter closes with a brief discussion of data gaps and research needs.

5.1. METHODS FOR ANALYSIS OF GENE EXPRESSION: DESCRIPTION OF MICROARRAY TECHNIQUES AND SEMI-QUANTITATIVE REVERSE TRANSCRIPTION-POLYMERASE CHAIN REACTION (RT-PCR)

5.1.1. Microarray Technology

Microarray is a technology that allows for simultaneous analysis of expression of thousands of genes from the organ or tissue of interest. In principle, there are two main types of microarrays: the cDNA microarray and the oligonucleotide array. The cDNA microarray contains DNA from each open reading frame spotted on to glass microscope slides or nylon membranes. These probes are used to detect cDNA, which is DNA synthesized from a mature, fully spliced mRNA transcript. For example, Clontech's Atlas Arrays contain DNA sequences from thousands of genes immobilized on nylon membrane or glass slides. Each gene found on these arrays is well characterized. These arrays, which use a radiolabelled detection system for analyzing the changes in gene expression, have been optimized for high-quality expression profiling using a limited set of genes. Moreover, they allow for the use of ^{32}P, and, therefore,

offer a sensitive measure of gene expression available. The second type of microarray is the oligonucleotide array. Here, short DNA sequences or oligonucleotides (oligos) are synthesized directly onto the glass slide via a number of different methods. For example, Affymetrix® uses a technique called 'Photolithographic' technology, wherein probes are directly synthesized on to the arrays. Briefly, the slide is coated with a light-sensitive chemical compound that prevents the formation of a bond between the slide and the first nucleotide of the DNA probe being created. Chromium masks are then used to either block or transmit light onto specific locations on the surface of the slide. A solution containing thymine, adenine, cytosine, or guanine is poured over the slide, and a chemical bond is formed in areas of the array that are not protected by the mask (exposed to light). This process is repeated 100 times in order to synthesize probes that are 25 nucleotides long. This method allows for high probe density on a slide.

Affymetrix® uses an antibody detection system with horseradish peroxidase and streptavidin conjugates, and a 2-dye system (Cy3- and Cy5- labeled fluorescein dyes), which is unique to this platform. The Agilent scanner detects the relative intensities of the red and green labels and gives a relative measure of the gene expression changes between the control and treated samples. In the case of Affymetrix® and Clontech, the detection system measures the absolute intensity of the individual probes of the treated and control samples. These values are then used to calculate the relative gene expression change between the treated and control samples.

5.1.2. Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

Polymerase Chain Reaction (PCR) is a method that allows exponential amplification of short DNA sequences within a longer double stranded DNA molecule using a thermo-stable DNA polymerase called Taq polymerase. RT-PCR is a semiquantitative technique for detection of expressed gene transcripts or mRNA. Over the last several years, the development of novel chemistries and instrumentation platforms enabling detection of PCR products on a real-time basis has led to widespread adoption of real-time RT-PCR as the method of choice for quantitating changes in gene expression. Real-time PCR is a kinetic approach in which the reaction is observed in the early, linear stages. Furthermore, real-time RT-PCR has become the preferred method for confirming results obtained from microarray analyses and other techniques that evaluate gene expression changes on a global scale.

5.2. REVIEW OF THE PUBLISHED DBP TOXICOGENOMIC STUDIES

5.2.1. Overview of the Toxicogenomic Studies

We evaluated nine studies published prior to July 2008 that characterized altered gene expression in rats following prenatal DBP exposure. Among these nine studies, four are based on the analysis of preselected genes by real-time RT-PCR, while the other five are based on the analysis of global gene expression by microarray technology. Table 5-1 summarizes general information (e.g., DBP dose, exposure route, exposure window, and tissue type) for these nine studies, and brief descriptions of each study are provided. Section 5.2.3.2 presents information about the similarities and differences among these studies.

5.2.2 Microarray studies

5.2.2.1. *Shultz et al. (2001)*

Six SD rats per group were treated by gavage with corn oil, DBP (500 mg/kg), or flutamide (reference antiandrogen, 50 mg/kg-d) from GD 12 - 16, GD 12 - 19, or GD 12 - 21. Testes were then isolated on GD 16, 19, or 21. Global changes in gene expression were determined by Clontech cDNA expression array (588 genes). Shultz et al. (2001) isolated total RNA from testis of control and treated animals. Reverse transcription reactions were performed using total RNA, [^{32}P]-dATP, and superscript II MMLV-RT. Following purification, the probes were counted, and equal numbers of counts per minute were added to each rat gene cDNA expression array. The arrays were hybridized with cDNA using 1 fetus per dam. Hybridization and washing were performed according to manufacturer's instructions. Digital images were collected on a BioRad phosphorimager and analyzed using Clontech's Atlas Image software. Eight genes were further examined by real-time RT-PCR. Total RNA was isolated from both testes using RNA STAT60, and then treated the RNA with DNase I with RNasin. cDNA was then synthesized using random primers and TaqMan reverse transcription reagents. Quality of RT reactions was confirmed by comparison of RT versus no enzyme control for each RNA sample using the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) primer set. Fourteen

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

Study[a]	Strain and species	DBP doses	Treatment interval[b]	Toxicogenomic method		Tissues collected
				Microarray (Platform)	RT-PCR	
Barlow et al., 2003	SD rat	500 mg/kg-d	GD 12−19	No	Yes	Testis
Bowman et al., 2005	SD rat	500 mg/kg-d	GD 12−19 or 19−21	Yes (Clontech cDNA arrays)	Yes	Wolffian ducts
Lehmann et al., 2004	SD rat	0.1, 1.0, 10, 50, 100, or 500 mg/kg-d	GD 12−19	No	Yes	Testis
Liu et al., 2005[c]	SD rat	500 mg/kg-d	GD 12−19	Yes (Affymetrix® GeneChip® oligo arrays)	Yes	Testis
Plummer et al., 2007[d]	Wistar rat	500 mg/kg-d	GD 12.5−15.5; 12.5−17.5, or 12.5−19.5	Yes (Agilent 22K and 44K oligo arrays)	Yes	Testis: whole, seminiferous cord, and interstitial regions
Shultz et al., 2001	SD rat	500 mg/kg-d	GD 12−16, 12−19, or 12−21	Yes (Clontech cDNA arrays)	Yes	Testis

Table 5. (continued)

Study[a]	Strain and species	DBP doses	Treatment interval[b]	Toxicogenomic method		Tissues collected
				Microarray (Platform)	RT-PCR	
Thompson et al., 2004	SD rat	500 mg/kg-d	GD 12−17, 18, or 19; 13−19, 14−19, 15−19, 16−19, 17−19, 18−19, or 19	No	Yes	Testis
Thompson et al., 2005	SD rat	500 mg/kg-d	0.5−24 hr on GD 18−19 or GD 19	Yes (Affymetrix® GeneChip® oligo arrays)	Yes	Testis
Wilson et al., 2004[e]	Rat, SD	1,000 mg/kg-d	GD 13−17	No	Yes	Testis

[a]In all studies, oral gavage was the route of exposure.
[b]GD 0 = sperm positive.
[c]Study assessed 7 different phthalates.
[d]Plummer et al. (2007) reported dosing intervals spanning GD 12.5−19.5, which is comparable to GD 12−19 in the other studies due to differences in reporting of GD and sperm positive at GD 0.5.
[e]Wilson et al. (2004) reported a dosing interval of GD 14−18, which is comparable to GD 13−17 in the other studies due to differences in reporting of GD and sperm positive at GD 1.

rat-specific primer sets were used for analyses. The ABI PRISM 7700 and the ABI PRISM 7900HT Sequence Detection System was used for RT-PCR, with the SYBR Green PCR and TaqMan Universal PCR Master Mix reagents. GAPDH was used as an on-plate internal calibrator for all RT-PCR reactions.

Genes analyzed by real-time RT-PCR include clusterin (*Clu*), cytochrome P450, family 11, subfamily a, polypeptide 1 (*Cyp11a1*), myristoylated alanine-rich C-kinase substrate (Marcks), proliferating cell nuclear antigen (*Pcna*), cytochrome P450, family 17, subfamily a, polypeptide 1 (*Cyp17a1*), steroidogenic acute regulatory protein (*Star*), scavenger receptor class B, member 1 (*Scarb1*), and v-kit Hardy Zuckerman 4 feline sarcoma viral oncogene homolog (*Kit*). Radioimmunoassay of steroid hormones and immunocytochemical analysis of certain proteins (i.e., CLU and b-cell leukemia/lymphoma 2 [BCL2]) in the fetal testes were also performed.

Of the 588 genes examined, ~45 genes had at least a 2-fold change in the average expression values in DBP-treated rats relative to the average values in control rats. DBP exposure led to a reduced expression of steroidogenic enzymes at GD 19, such as *Cyp11a1, Cyp17a1, Scarb1*, and *Star*. These genes were upregulated at GD 19 following flutamide exposure, suggesting that DBP does not act as an androgen antagonist at this time point. Flutamide and DBP demonstrate patterns of gene expression that overlap, though both have distinctly expressed genes. This suggests to Shultz et al. (2001) that there are both common and distinct molecular pathways within the developing fetal testes.

Other genes affected after DBP exposure were *Clu* (upregulated) and *Kit* (downregulated). Using immunocytochemical staining of CLU and BCL2 protein in the fetal testes, increased amounts of both proteins were observed in the Leydig and Sertoli cells of GD 21 testes. Decreases in testicular T and androstenedione in testes isolated on GD 19 and 21 were observed, while increases in progesterone in testes isolated on GD 19 in DBP-exposed testis were observed.

Shultz et al. (2001) suggest that the antiandrogenic effects of DBP are due to decreased T synthesis. Furthermore, enhanced expression of cell survival proteins, such as CLU and BCL2, may be involved in DBP-induced LC hyperplasia, while downregulation of c-KIT may play a role in gonocyte degeneration.

5.2.2.2. *Bowman et al. (2005)*

Four to seven SD rats per group were treated by gavage with corn oil or DBP at 500 mg/kg-d from GD 12 to 19 or GD 12 to 21. The animals were sacrificed on GD 19 or 21, and Wolffian ducts (WD) were pooled from three to four fetuses (to obtain enough RNA for analysis) within the same litter for gene expression analysis. Global changes in gene expression were determined by Clontech Atlas Rat Toxicology 1.2 cDNA expression array (1,185 genes). Images were collected using a PhorpshorImager and then imported into AtlasImage 2.01 and GeneSpring 4.2 for analysis. Selected genes were further examined by real-time quantitative RT-PCR using the GeneAmp 5700 Sequence Detection System. Total RNA was isolated, DNAse-treated, and reverse-transcribed using TaqMan reagents. Twenty-three primer sets were used for RT-PCR analysis. Reactions were standardized using GAPDH-specific primers. The genes analyzed by RT-PCR include those in the insulin-like growth factor (Igf) pathway, the matrix metalloproteinase (Mmp) family, the extracellular matrix, and in other developmentally conserved signaling pathways: bone morphogenetic protein 4 (*Bmp4*), collagen, delta like (*Map3k12*), epidermal growth factor receptor (*Egfr*), fibroblast growth factor 10 (*Fgf10*), FGF receptor 2 (*Fgfr2*), fibronectin, insulin-like growth factor 1 (*Igf1*), insulin-like growth factor 2 (*Igf2*), insulin-like growth factor 1 receptor (*Igf1r*), insulin-like growth factor binding protein 5(*Igfbp5*), integrinA5, integrinB1, matrix Gla protein (*Mgp*), matrix metalopeptidase 2 (*Mmp2*), matrix metalopeptidase 14 (*Mmp14*), matrix metalopeptidase 16 (*Mmp16*), Notch2 receptor (*Notch2*), and tissue inhibitors of MMPs *(Timp1, Timp2,* and *Timp3)*. Immunohistochemistry was also performed to evaluate changes in localization and/or intensity of IGFLRβ and androgen receptor (AR) protein expression.

Microarray data were not presented due to considerable variability in gene expression levels within the treatment group at each age. Based on real-time PCR analysis, compared with controls, prenatal exposure to DBP from GD 12 to 19 or GD 12 to 21 increased mRNA expression of different members of the IGF family including *Igf1* (on GD 19 and 21)*, Igf 2*(on GD 19)*, Igfr1r* (on GD 19)*,* and *Igfbp5* (on GD 21) in the developing WD, while *Egfr* was unchanged on GD 19 and GD 21. Additionally, mRNA expression of *Ar, Bmp4, integrinA5, Mmp2,* and *Map3k12* was increased on GD 19; mRNA expression of *Fgf10, Fgfr2, Notch2, Mmp2, Timp1,* and *Mgp* was increased on GD 21. IGFLRβ immunostaining was higher in the cytoplasm of the ductal epithelial cells and increased in the cytoplasm of mesenchymal cells in

DBP-exposed fetuses compared with that in controls. In general, reduction of AR immunostaining in the nuclei of ductal epithelial cells of DBP-exposed WD was observed on GD 19. Compared with controls, WDs dissected from GD 19 DBP-exposed fetuses were slightly smaller in size (underdeveloped) and appeared to be more fragile. By GD 21, control fetus WDs were markedly coiled, while those from the exposed fetuses exhibited less coiling.

Prenatal DBP exposure appears to alter the mesenchyme-epithelial signaling of growth factors (e.g., IGFs) and other developmentally conserved pathways (e.g., BMP4) in WDs. Bowman et al. (2005) contend that the effect of DBP on WD differentiation is likely a consequence of decreased fetal testicular T, although direct effects of DBP on the developing WD independent of T are also possible.

5.2.2.3. *Liu et al. (2005)*

Five to ten SD rats per group were treated by gavage with corn oil, DBP (500 mg/kg-d), or one of six other phthalate esters (500 mg/kg-d) daily from GD 12 to 19. The six other phthalate esters include diethyl phthalate (DEP), dimethyl phthalate (DMP), diocytyl tere-phthalate (DOTP), diethylhexyl phthalate (DEHP), dipentyl phthalate (DPP), and butyl benzyl phthalate (BBP). Testes were collected on GD 19, homogenized, and then, total RNA was isolated. RNA integrity was assessed using an Agilent 2100 Bioanalyzer. cDNA was synthesized from 2.5 μg total RNA, and purified using RiboAmp OA. The BioArray High-Yield RNA Transcript Labeling Kit was used for cRNA amplification and biotin labeling. Affymetrix® GeneChip Sample Cleanup Module was used for purifying and fragmenting the cRNA. The Complete GeneChip® Instrument System was then used to hybridize, wash, stain, and scan the GeneChip® arrays (RAE230A and RAE230B; ~30,000 genes). The data were analyzed using analysis of variance (ANOVA [one-way, two-way, nested one-way]), Dunnett test (post hoc), Tukey test, and Bonferroni adjustment.

Image files obtained from the scanner were analyzed with the Affymetrix® Microarray Suite (MAS) 5.0 software and normalized by global scaling. Absolute analysis was performed for each array prior to comparative analysis. To identify differentially expressed transcripts, pair-wise comparison analyses were carried out with MAS 5.0 (Affymetrix®). P-values were determined by the Wilcoxon's signed rank test and denoted as "increase", "decrease", or "no change". A transcript is considered significantly altered in relative abundance when $p < 0.05$.

Analysis using MAS 5.0 provides a signal log ratio (SLR), which estimates the magnitude and direction of change of a transcript when two arrays are compared (experimental versus control). The SLR output was converted into "fold-change" as recommended by Affymetrix®. Furthermore, stringent criteria were used to identify robust signals as follows: (1) software call of "present", (2) ≥2.0-fold change or SLR 1.0, in both replicates. Average and standard deviations were calculated for all the fold-change values. In general, only transcripts induced or suppressed by ≥2-fold were considered as differentially expressed.

Selected genes were further examined by real-time quantitative RT-PCR using 18 primer sets. The genes analyzed by RT-PCR include epididymal secretory protein 1 (*re1*), low-density lipoprotein receptor (*Ldlr*), 17β-hydroxysteroid dehydrogenase 3 (*Hsd17b3*), 17β-hydroxysteroid dehydrogenase 7 (*Hsd17b7*), luteinizing hormone/choriogonadotropin receptor (*Lhcgr*), CCAAT/enhancer-binding protein (C/EBP), beta (*Cebpb*), early growth response 1 (*Egr1*), nuclear receptor subfamily 4, group A, member 1 (*Nr4a1*), nuclear factor, interleukin 3, regulated (*Nfil3*), nuclear receptor subfamily 0, group B, member 1 (*Nr0b1*), transcription factor 1 (*Tcf1*), insulin-induced gene 1 (*Insig1*), protein kinase C-binding protein (*Prkcbp1*), decay-accelerating factor (*Daf*), dopa decarboxylase (*Ddc*), seminal vesicle secretion 5 (*Svs5*), and testis-derived transcript (*Testin*). AGD was measured and immunohistochemistry was performed for NR0B1, TESTIN, GEB14, DDC, and CEBPB proteins.

Of ~30,000 genes examined, 391 were statistically significantly altered following exposure to the four developmentally toxic phthalates (DBP, BBP, DPP, and DEHP) relative to the controls. While the four developmentally toxic phthalates were indistinguishable in their effects on global gene expression, no significant changes in gene expression were detected in the phthalates that do not lead to developmental effects (DMP, DEP, and DOTP). Of the 391 genes altered by the developmentally toxic phthalates, 225 were unknown and uncharacterized transcribed sequences. Of the remaining 166 genes, the largest GO classification (31 genes) was of genes related to lipid, sterol, and cholesterol homeostasis. Additional GO classification groups include genes involved in lipid, sterol, and cholesterol transport (10 genes); steroidogenesis (12 genes); transcription factors (9 genes); signal transduction (22 genes); oxidative stress (11 genes); and cytoskeleton-related (13 genes). RT-PCR results indicated that the developmentally toxic phthalates reduced the mRNA levels of *Hsd17b7, Lhcgr, Ldlr, re1, Svs5, Insig1,* and *Ddc.* Additionally, the RT-PCR results indicated that the developmentally

toxic phthalates induced the mRNA levels of *Grb14, Prkcbp1*, and *Testin*. RT-PCR results also indicated that gene expression of several transcription factors including *Dax-1, Cebpb, Nfil3, Nr4a1,* and *Tcf1* were significantly changed by at least one of the toxic phthalates. Based on immunohistochemical analysis, DAX-1 expression was reduced in the gonocyte population of DBP-treated testis compared with that of controls. Additionally, the expression of nuclear CEBPB, GRB14, and DDC proteins was reduced in interstitial cells of DBP treated testis, while TESTIN and GRB14 expression levels were increased in Sertoli cells of DBP treated testis. An AGD reduction was observed in male fetuses exposed to any of the developmentally toxic phthalates.

This study showed that the four phthalates (DBP, DEHP, BBP, and DPP) that have similar effects on the developing male rat reproductive tract are indistinguishable in their genomic signature for the developing fetal testis. These phthalates targeted pathways in Leydig cell production of T and other pathways that are important for normal interaction and development between Sertoli cells and gonocytes. By contrast, in animals exposed to any of the four phthalates that have not exhibited developmental toxicity (the "nondevelopment" phthalates) did not have the same genomic signature.

5.2.2.4. *Thompson et al. (2005)*

Four SD rats per group were gavaged with corn oil or DBP at 500 mg/kg-d daily. In the first study, the treatment was performed on GD 18 or GD 19, followed by animal sacrifice 30 min, 1 hr, 2 hr, 3 hr, 6 hr, 12 hr, 18 hr, or 24 hr after the treatment on GD 19. Global changes in gene expression were determined by Affymetrix® GeneChips® (GeneChips® used in the study were not reported). The methods were similar to Liu et al. (2005)—with the exception of the statistical analysis. Thompson et al. (2005) used JMP statistical software to perform Student t-tests or one-way ANOVAs with Tukey post hoc analysis. Selected genes were further examined by real-time quantitative RT-PCR. An ABI Prism 7900HT Detection System, the SYBR Green PCR Master Mix, and 30 primer pairs were used for analysis of DBP-induced changes in gene expression. The genes analyzed by RT-PCR included *Cyp11a1, Scarb1, Star, Cyp17a1, Egr1, Egr2, Nr4a1, Nfil3, Tcf1,* serum/glucocorticoid regulated kinase (*Sgk*), tumor necrosis factor receptor superfamily, member 12a (*Tnfrsf12a*), sclerostin domain containing 1 (*Sostdc1*), Wnt oncogene homolog 4 (*Wnt4*), B-cell translocation gene 2, antiproliferative (*Btg2*),

C/EBP, delta (*Cebpd*), FBJ murine osteosarcoma viral oncogene homolog (*Fos*), dual specificity phosphatase 6 (*Dusp6*), *Hes6_predicted*, interferon-regulated developmental regulator (*Ifrd1*), *Ldlr*, nuclear receptor subfamily 4, group A, member 3 (*Nr4a3*), *Pawr*, *Nr0b1*, Jun-B oncogene (*Junb*), endothelial differentiation sphingolipid G-protein-coupled receptor 3 (*Edg3*), thrombospondin 1 (*Tsp1*), and stanniocalcin 1 (*Stc1*). Immunoblotting by SDS-PAGE was performed for SCARB1, CYP11a1, STAR, and CYP17a1. Fetal testicular T concentration was determined by radioimmunoassay.

Based on microarray analysis, there were 106 genes in the DBP-treated groups that were significantly different from time-matched controls. Six genes were significantly elevated within 1 hour of DBP exposure. An additional 43 genes were upregulated, and 5 genes were downregulated 3 hours after DBP exposure. The rapid induction of these genes was a transient effect; none of the genes upregulated 1 hour after DBP treatment were still significantly different than the control group 6 hours after treatment. Only nine genes showed significant changes from the control group between the 3- and 6-hour time points. Prior to 6 hours after DBP exposure, the majority of the changes in expression had reflected increased transcription. At 6 hours, 19 genes were significantly decreased, and 17 were increased in expression. Based on RT-PCR analysis, the immediate early gene *Fos* and the putative mRNA destabilizing gene zinc finger protein 36 (*Zfp36*) were at peak expression level 1 hour after DBP exposure. Other immediate early genes were at peak expression at 2 hours after DBP exposure. At 3 hours after exposure, the expression of *Cebpd*, *Cxcl1*, and *Nr4a3* increased rapidly, while other genes showed a more gradual increase. *Tsp1* expression was increased 25-fold at 3 hours and returned to baseline at 6 hours. Genes involved in testicular steroidogenesis were first noticeably affected 2 hours after DBP exposure: Inhibition of *Star* transcription was detected ~2 hours after DBP exposure. *Scarb1, Cyp11a1,* and *Cyp17a1* showed a significant decrease in expression at about 6 hours after DBP exposure. Also, after 6 hours, the T concentration dropped to approximately the level observed after long term DBP treatment. After 12 hours of exposure, steroidogenesis-associated genes *Nr0b1* and *Nr4a1* were elevated. *Tcf1* and *Sgk* were downregulated soon after DBP exposure, but values returned to control levels by 3 hours after DBP exposure. *Sostdc1* and *Hes6_predicted* returned to control levels at 6 hours after exposure. Based on radioimmunoassay, a decrease in fetal testicular T to 50% was observed within an hour after DBP exposure. In a second experiment to compare the effect of DBP on steroidogenesis in the

fetal adrenal gland, DBP treatment at GD 12−19 was followed by analysis of gene expression in this tissue. A decrease (but not statistically significant) of corticosterone after GD 12−19 DBP exposure was observed in the fetal adrenal. The expression of genes involved in steroidogenesis was less affected in the adrenal (males and females) than in the testes. This study indicates that the effect of DBP exposure on steroidogenesis gene expression is specific to the fetal testis and not in other steroidogenic organs.

Rapid transcriptional changes after DBP exposure in a number of genes could be responsible for the reduction in steroidogenesis. Peroxisome proliferator-activated receptors (PPAR) activation is ruled out since changes in expression of genes targeted by PPAR α and γ are not observed until 3 hours after DBP treatment. Many of the genes whose upregulation was detected within the first hour after treatment were "immediate early genes," meaning genes involved in cell growth and differentiation. One possible mechanism for DBP's repression of steroidogenesis is that DBP may initially stimulate the mitogen-activated protein kinase/extracellular signal-regulated kinase (MAPK/ERK) pathway in the fetal testis. Increased expression of *Egr1* and *Zfp36* could, in turn, lead to degradation of the transcripts involved in testicular steroidogenesis. Consistent with this possibility, the *Star* mRNA contains the AU-rich element, which are regions with many A and U bases that target the RNA for degradation, in target transcripts of *Zfp36*.

5.2.2.5. *Plummer et al. (2007)*

Five Wistar rats per group were gavaged with corn oil or DBP at 500 mg/kg-d from GD 12 until the day prior to sacrifice. Animals were sacrificed on GD 15, 17, or 19 and used for immunolocalization, Western analysis, or RNA quantification (of whole testes, seminiferous cord, or interstitial regions using laser capture microdetection). Samples for laser capture microdetection were collected from sections of single testes from GD 19 animals. RNA samples from three treated litters were compared to a pool of RNA samples from control animals to lessen errors due to biological variation. The Agilent 22K rat and 44K whole-rat oligonucleotide arrays were used for analysis of the whole-fetal testes and microdissected tissue, respectively. RNA was isolated from the homogenized whole-fetal testes using the RNeasy mini kit (Qiagen) and from laser capture microdissected samples using RNeasy micro kit (Qiagen). Isolated RNA was labeled using the Agilent Low Input Linear Amplification Labelling kit according to the

manufacturer's instructions. Specific activity of the labeled cRNA was measured using the microarray analysis program on a NanoDrop ND1000 spectrophotometer (Montchanin, USA). Microarray analysis with whole-fetal testis RNA was performed using Agilent 22K rat oligonucleotide arrays (Agilent #G4110A). Regional microarray analysis on RNA isolated from laser capture microdissected fetal testis tissue was performed using Agilent 44K whole-rat genome oligonucleotide microarrays (Agilent #G4131A). Microarray data analysis was conducted using Agilent feature extraction (v7.1) and Rosetta Luminator software (Rosetta Biosoftware, Kirkland, USA) to generate "signature" lists, defined as significantly ($p < 0.01$) different. The compare biosets function in Luminator was used to compare signature lists from different fetal testis regions. Pathway analysis used Ingenuity Pathways Analysis software.

DBP induced statistically significant changes in gene expression at all three time points. At GD 15 in whole testes, expression of genes regulating lipid metabolism, redox homeostasis, cell proliferation, and apoptosis were altered. At GD 17 and 19, these four main gene clusters were altered: steroidogenesis (e.g., *Cyp17a1, Cyp11a1*), lipid metabolism, cholesterol (e.g., *Star, Scarb1*), and redox homeostasis. In laser- capture microdissection studies of GD 19 tissue, both regions demonstrated altered expression of genes associated with steroidogenesis (e.g., *Cyp17a1*), cholesterol transport (e.g., *Scarb1*), cell/tissue assembly, and cellular metabolism. In the interstitial regions only, genes involved in fatty acid oxidation, testes morphogenesis, and descent (e.g., *Insl3*) were altered. In the cord samples, gene associated with stress responses, chromatin bending, and phagocytosis were altered.

RT-PCR analysis was performed on RNA from GD 19 testes from five rats/group using sequence specific primers for the orphan nuclear receptor, steroidogenic factor 1 (*Sf-1), Star, Cyp11a,* and *Insl3*. The data were analyzed using a one-way ANOVA followed by the Bonferroni post-test, using GraphPad Prism. These studies showed a statistically significant reduction in the expression of *Star, Cyp11a1,* and *Insl3* but not *Sf-1*.

Analysis of protein expression at GD 19 showed DBP-induced reduction in levels of CYP11A, inhibin-α, cellular retinoic acid binding protein 2 (CRABP2), and phosphatidylethanolamine binding protein (PEBP) in Leydig cells, and no change in Sertoli cells/seminiferous cords. These data correlated with microarray data for the genes coding for these proteins. Immunoreactivity for antimullerian hormone (AMH) was slightly increased in Sertoli cells following DBP treatment. Western blot analysis and immunolocalization of SF1

demonstrated no effects of DBP on protein expression in Sertoli or LCs. Using time plots to assess time-dependent changes in gene expression, a coordinate down-regulation of inhibin-α, Scarb1, Star, and Cyp11a1A1 was observed between GD 15 and 19.

This study confirms other study results, showing down-regulation of *Scarb1, Star, Cyp11a1,* and *Cyp17a1*. The authors suggest that DBP induces LC dysfunction indirectly through sequestration of cofactors used in key signaling pathway and not through decreases in SF1 protein expression. They further state that the use of Wistar rats could be important, as Wistar rats may be more susceptible than SD rats to testicular effects of DBP.

5.2.3. Real-Time Reverse Transcription-Polymerase Chain Reaction (RT-PCR) Studies

5.2.3.1. *Barlow et al. (2003)*

Six to seven SD rats per group were gavaged with corn oil or DBP at 500 mg/kg-d from GD 12 to 19. Testicular RNA was then isolated from three randomly selected male fetuses per litter. RT-PCR studies were performed as described in Shultz et al. (2001).

mRNA of 13 preselected genes in the steroid biosynthetic pathway was analyzed by real-time RT-PCR; immunohistochemical and oil red O histochemical analyses were performed to further confirm mRNA changes. The 13 genes analyzed were *Scarb1, Star, Cyp11a1,* hydroxyl-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1 *(Hsd3b), Cyp17a1,* hydroxysteroid (17-beta) dehydrogenase 3 *(Hsd17b3), Ar,* luteinizing hormone receptor *(Lhr),* follicle-stimulating hormone receptor *(Fshr), Kit,* stem cell factor *(Scf), Pcna,* and *Clu.*

Compared with controls, mRNA expression was downregulated for *Scarb1, Star, Cyp11a1, Hsd3b, Cyp17a1,* and *Kit* in DBP-treated testes; mRNA expression was upregulated for *Clu* following DBP exposure. These changes in mRNA expression were supported by immunohistochemical localization of selected proteins and by staining for lipids.

The results in the study of Barlow et al. (2003) confirm the gene expression changes observed in a previous study (Shultz et al., 2001). Furthermore, the data support alterations in cholesterol synthesis, transport, and storage that likely play a role in decreased T production by fetal LCs. The decreased level of mRNA expression for *P450scc* indicates another possible contributor, as P450scc conversion of cholesterol to pregnenolone is the limiting enzymatic step in T biosynthesis.

5.2.3.2. *Lehmann et al. (2004)*

To date, Lehmann et al. (2004) is the only dose-response gene expression study on the testis performed with DBP. The other studies used a single high dose shown to affect male reproductive system development.

Five to seven SD rats per group were treated by gavage with corn oil or DBP at 0.1, 1.0, 10, 50, 100, or 500 mg/kg-d from GD 12−19. Testes were then isolated on GD 19, and changes in gene and protein expression were measured by real-time RT-PCR (as described in Shultz et al., 2001) and Western analysis. Ten preselected genes in the steroid biosynthetic pathway were analyzed by RT-PCR: *Scarb, Star, Cyp11a1, Hsd3b1, Cyp17a1, Kit,* benzodiazepine receptor, peripheral *(Bzrp),* insulin-like 3 *(Insl3), Clu,* and sterol regulatory element binding factor 1 (*Srebf1*). Fetal testicular T concentration was determined by radioimmunoassay in a separate group of animals using doses of 0.1, 1.0, 10, 30, 50, 100, or 500 mg/kg-d.

The aim of this study was to determine the DBP doses at which statistically significant alterations in the expression of a subset of genes and a reduction in fetal testicular T occur. As summarized in Table 5-2, Lehmann et al. (2004) established 50 mg DBP/kg-d as a LOEL and 10 mg DBP/kg-d as a NOEL for reductions in genes and proteins associated with T production as well as genes associated with other MOAs (e.g., *Kit, Insl3*) together with reductions in intratesticular T. The Lehmann et al. (2004) study demonstrated that *Hsd3b (*also called *3β-HSD)* gene expression involved in T synthesis was detected at levels as low as 0.1 mg/kg-d.

DBP exposure resulted in a dose-dependent decline in expression of the genes involved in cholesterol transport and steroidogenesis: *Scarb1, Star, Cyp11a1, Hsd3b, Cyp17a1,* and *Insl3*. Expression of *Bzrp* and *Clu* were increased in response to DBP. Furthermore, fetal testicular T was significantly reduced at DBP doses ≥50 mg/kg-d and reduced by 26% at 30 mg/kg-d. This study reported a LOEL of 50 mg DBP/kg-d and a NOEL of 10 mg DBP/kg-d for reductions in genes and proteins associated with T production together with reductions in intratesticular T. It demonstrates the coordinate reduction in genes and corresponding proteins involved in steroidogenesis and cholesterol transport, concurrent with a decrease in intratesticular T. Importantly, it shows effects on the male reproductive system at lower doses than are used in the other DBP studies in this review.

Table 5-2. Lehmann et al. (2004) dose-response gene expression change data[1] measured by RT-PCR showing statistically significant changes ($p < 0.05$).

Gene Symbol (reported gene name)	Dose (mg/kg-d)					
	0.1	**1**	**10**	**50**	**100**	**500**
Scarb1 (Sr-B1)	---[2]	↓0.6	---	↓0.5	↓0.3	↓0.2
Star	---	---	---	↓0.4	↓0.3	↓0.1
Cyp11a1 (P450ssc)	---	---	---	↓0.6	↓0.7	↓0.2
Cyp17a1	---	---	---	---	---	↓0.3
Hsd3b (3β-HSD)	↓0.3	↓0.4	---	↓0.5	↓0.3	↓0.5
Bzrp (PBR)	---	---	---	---	---	↑2.0
Trpm2	---	---	---	---	---	↑1.6
Kit (c-Kit)	↓0.3	↓0.5	---	↓0.3	↓0.5	↓0.1
Insl3	---	---	---	---	---	↓0.3

[1]Gene expression values are from DBP-exposed testes expressed relative to control values. They are the statistically significant averages from five separate rat fetuses from different dams per treatment group.
[2] --- = no statistically significant change.

Estimates for human exposure to DBP range from 0.84 to 113 μg/kg-d (0.00084 to 0.113 mg/kg-d). For *Scarb1, Hsd3b,* and *Kit,* significant reductions in mRNA levels were observed at DBP doses that approach maximal human exposure levels, 0.1 mg/kg-d. Alterations in the expression of *Scarb1, Hsd3b,* and *Kit* may be sensitive indicators of DBP exposure, but they are not necessarily of adverse consequences to DBP.

In another dose response study, Mylchreest et al. (2000) exposed pregnant SD rats to 0-, 0.5-, 5.0-, 50-, 100-, or 500-mg/kg-d DBP from GD 12−21. They found hypospadias and absent or partially developed ventral prostate, seminal vesicles, vas deferens, and epididymis at the 500 mg/kg-d dose. They reported a NOAEL and LOAEL of 50 and 100 mg/kg-d, respectively.

5.2.3.3. *Thompson et al. (2004)*

Four to five SD rats per group were gavaged with corn oil or DBP at 500 mg/kg-d from GD 12−19. Testes were isolated on GD 17, 18, or 19. Testes mRNA was isolated and four preselected genes (*Scarbl, Star, Cyp11a1,* and *Cyp17a1*) in cholesterol and steroidogenesis pathways was analyzed by real-time RT-PCR as described in Shultz et al. (2001). Immunoblotting was performed using total protein extracted from paired testis, and quantification of the expressed protein levels was done using FluorChem. Fetal testicular T concentration was determined by radioimmunoassay, and whole-cell cholesterol uptake assessment was performed on overnight cultures.

A significant decrease in fetal testicular T concentration was observed as early as GD 17 after in utero exposure of fetuses to DBP. The percent difference in testicular T between control and treated testes was much higher on GD 18 (17.8% of that seen in the control samples) than on GD 17 (46.6%). Furthermore, significant decreases in mRNA expression of *Scarbl, Star, Cyp11a1,* and *Cyp17a1* were observed as early as GD 17. In agreement with T levels, the percentage difference of gene expression between control and treated testes was higher on GD 18 than on GD 17. The suppression of the transcription by DBP was a reversible effect, as the mRNA levels for all the genes returned to control levels 48 hr after DBP withdrawal. When protein expression was analyzed, results similar to the gene expression data were obtained (i.e., strong expression in controls, decreased expression in treated animals with 24-hr DBP withdrawal, and rising expression after the 48-hr DBP withdrawal). Additionally, there was a significant decrease in the amount of cholesterol transported across the mitochondrial membrane in the testes from DBP treated fetuses as assayed in overnight cultures of testis explants. This observation indicates that the decrease in *Star* mRNA correlated with diminished protein function (transport of cholesterol from the outer to the inner mitochondrial membrane by the StAR protein is one of the rate-limiting steps of steroidogenesis).

The results of this study demonstrate that DBP-induced suppression of T production in the fetal testis correlate with diminished transcription of several genes in the cholesterol transport and steroidogenesis pathways as early as GD 17. This diminished effect was reversible, suggesting that DBP directly interferes with the signaling processes necessary for maintenance of steroidogenesis or with the transcriptional regulators required to maintain coordinate expression of the genes involved in cholesterol transport and T biosynthesis.

5.2.3.4. *Wilson et al. (2004)*

In the studies of Wilson et al. (2004) three to five SD rats per group were treated by gavage with corn oil or a developmental toxicant daily from GD 14−18 for two separate experiments. In the first experiment, the rats were treated with DEHP at 750 mg/kg-d. In the second experiment, the rats were treated with one of six chemicals, each known to induce male reproductive malformations. The chemicals used for the second study were three AR antagonists (vinclozolin [200 mg/kg-d], linuron [100 mg/kg-d], and prochloraz [250 mg/kg-d]) and three phthalate esters (DEHP [1 g/kg-d], DBP [1 g/kg-d], and BBP [1 g/kg-d]). Dams were killed on GD 18, and testes were removed and pooled by litter. In the first study, RNA was prepared to quantify expression of one preselected gene, *Insl3*, by real-time RT-PCR. In the second study, both steroid hormone production (ex vivo incubation) and *Insl3* expression were assessed. Total RNA was isolated using Trizol, digested using Dnase I, and quantitated with RiboGreen. ImProm-II Reverse Transcriptase was used for RT, followed by amplification using Taq1. They completed RT-PCR for *Insl3* using a Bio-Rad iCycler.

In the first study, the mRNA expression of *Insl3* was reduced by ~80% in DEHP litters compared with that in control litters. In the second study, among the six chemicals tested, only phthalate esters (DEHP, DBP, or BBP) reduced mRNA levels in the fetal testis, with DBP and BBP being more effective than DEHP. In contrast, prochloraz or linuron as well as any of the three phthalate esters significantly reduced ex vivo T production.

In a previous study with antiandrogenic chemicals that alter male sexual differentiation (Gray, et al. 2000), phthalate esters were the only class that produced agenesis of the gubernacular ligaments; some of the phthalate ester-exposed rats had a phenotype similar to that seen in the *Insl3* knock-out mouse. The study of Wilson et al. (2004) confirms this hypothesis since only the three phthalates reduced *Insl3* gene expression. The authors proposed that the effects of DEHP, DBP, or BBP on *Insl3* mRNA and T production result from a delay in maturation of fetal LCs, resulting in hyperplasia as they continue to proliferate rather than to differentiate.

5.2.4. Study Comparisons

5.2.4.1. *Microarray Study Methods Comparison*

Table 5-3 compares the study design and method of determining statistical significance across the five microarray studies used in the case study. Because the Bowman et al. (2005) paper assessed changes in gene expression in WD rather than testis, and because the microarray data were not presented in the paper, the discussions will focus on the three other microarray studies. The Plummer et al. (2007) study pooled control tissue and used the Agilent platform, which differed from the platforms used in the other studies. Liu et al. (2005), Schutz et al. (2001), and Thompson et al. (2005) all assessed mRNA levels in rat testis—but with somewhat differing significance criteria. All studies included vehicle-treated controls.

5.2.4.2. *Reverse Transcription-Polymerase Chain Reaction (RT-PCR) Study Methods Comparison*

Table 5-4 compares the RT-PCR methods used in the nine toxicogenomic articles. There were many similarities among the studies. All the groups—except Bowman et al. (2005)—extracted RNA from testis. All studies used a vehicle-treated control. Most of the studies used the same significance criteria ($p < 0.05$). There were some differences in the number of fetuses used per experiment while some studies pooled tissues.

More important, however, were the significant similarities among the nine toxicogenomic studies. Eight of the studies used the same strain of rat (SD), all purchased from the same vendor (Charles River, Raleigh, NC). All of the studies described dissolving the DBP in corn oil, using a corn oil vehicle control, and using oral gavage as the route of exposure. Six of the studies (Barlow et al., 2003; Bowman et al., 2005; Liu et al., 2005; Shultz et al., 2001; Plummer et al., 2007; Thompson et al., 2004) treated the animals by gavage to 500 mg/kg-d from GD 12−19. This dose has been shown to adversely affect male reproductive development without causing maternal toxicity or fetal death. Lehmann et al. (2004) completed a dose response during the GD 12−19 period, using 0.1, 1.0, 10, 50, 100, or 500 mg/kg-d. Bowman et al. (2005) and Shultz et al. (2001) included an additional exposure duration of GD 12−21. Wilson et al. (2004) exposed for a slightly shorter duration (GD 13−17) and at a higher dose (1000 mg/kg-d). This paper reports exposures on GD 14−18; however, these authors consider GD 1 as the day a sperm-positive smear was identified in dams, whereas the other studies consider the

Table 5-3. Method comparisons for DBP microarray studies

Study	Tissue collected	Significance criteria	Individual animals (*n*) used?
Bowman et al., 2005	Wolffian ducts	ND (microarray data not presented)	No, pooled (3−4 fetuses/litter; 67 dams/treatment group)
Liu et al., 2005	Testis	$p < 0.05$ compared to control by either 1-way ANOVA, post hoc Dunnett test, or Tukey test	Yes, (6 fetuses/litter; 6 dams/treatment group)
Plummer et al., 2007	Testis (whole, laser captured interstitial tissue, or laser captured seminiferous cord tissue)	$p < 0.01$ using Agilent feature extraction software and then Rosetta Luminator software by performing one-way ANOVA on log fold change in the replicates	Yes for DBP-treated (3 pups from 3 different dams); Control RNAs were pooled
Shultz et al., 2001	Testis	2-fold change in average expression value compared to control	GD 19 and GD 21 time points: Yes, 1 fetus/litter; 3 dams/treatment group. GD 16 timepoint: pooled RNA from 5 fetuses/1 litter; 3 arrays hybridized/treatment group.
Thompson et al., 2005	Testis	$p < 0.05$ multiple comparison using Bonferroni correction	Yes (ND)

ANOVA, analysis of variance; ND, not detected.

Table 5-4. Method comparisons among the reverse transcription-polymerase chain reaction (RT-PCR) DBP studies

Study	Tissue collected	Significance criteria (*p* values)	Individual animals (*n*) used?
Barlow et al., 2003	Testis	$p < 0.05$ compared to control	Yes (3 fetuses/litter; 5 dams/treatment group)
Bowman et al., 2005	Wolffian ducts	$p < 0.05$ compared to control	No, pooled (3−4 fetuses/litter; 6−7 dams/treatment group)
Lehmann et al., 2004	Testis	$p < 0.05$ compared to control	Yes (5 fetuses/litter; 4−5 litters/treatment group)
Liu et al., 2005	Testis	$p < 0.05$ compared to control by either 2-way nested ANOVA or Dunnett	Yes (control: 6 fetuses/dam; 6 dams for control. Treated: 3 fetuses/dam; 3 dams)
Plummer et al., 2007	Testis (whole, laser-captured interstitial tissue, or laser-captured seminiferous cord tissue)	$p < 0.05$ compared to control, normalized to 1.0. Expressed as mean + / – SEM; one-way ANOVA followed by Bonferroni post test using GraphPad Prism software	ND[a]; assessed GD 19.5 fetal testes
Shultz et al., 2001	Testis	$p < 0.05$ compared to control	GD 19 and GD 21 timepoints: Yes, 1 fetus/litter; 3 dams/treatment group. GD 16 timepoint: pooled RNA from 5 fetuses/1 litter; 3 arrays hybridized/treatment group.
Thompson et al., 2004	Testis	$p < 0.05$ compared to control (Student's t-test or 1-way ANOVA)	ND
Thompson et al., 2005	Testis	$p < 0.05$ normalized mean of 3−5 fetuses/treatment group relative to control	Yes, 3−5 fetuses/litter
Wilson et al., 2004	Testis	$p < 0.01$ compared to control (means on a litter basis)	No, pooled for each litter (3 dams/treatment group)

[a]Not clear from the Materials and Methods.

ANOVA, analysis of variance; ND, not detected

sperm-positive day as GD 0. Therefore, to be consistent with the other groups, we are reporting the exposure period as GD 13−17. Similarly, Plummer et al. (2007) reports exposures ranging from GD 12.5 to GD 19.5, which are equivalent to GD 12−19 as the authors consider GD 0.5 to be the sperm positive day, have been adapted, too.

All of the other selected studies collected testes for RNA extraction, with the exception of Bowman et al. (2005), which collected WDs. Bowman et al. (2005) focused on the WD because they were interested in characterizing the mechanisms responsible for prenatal DBP-induced epididymal malformations. WD tissue from 3−4 fetuses was obtained to ensure enough RNA for analyses (Table 5-3). Since WDs are the precursor of the vas deferens, epididymis, and seminal vesicles, the tissue assayed by Bowman et al. (2005) is different from the tissue evaluated in the other seven studies (RNA from the testes of 1−3 fetuses). The studies used a variety of toxicogenomic methodologies to assess changes in gene expression. General descriptions of these methods utilized by the studies were presented in Section 5.1.

An important consideration is the reliability of the data being generated and compared in these nine DBP studies. As discussed, the MAQC project (MAQC Consortium et al., 2006) has recently completed a large study evaluating inter- and intraplatform reproducibility of gene expression measurements (see Chapter 2). Six commercially available microarray platforms and three alternative gene expression platforms were tested. Both Affymetrix® microarrays and RT-PCR assays were included in the MAQC testing. Affymetrix® and the other one color platforms showed similar coefficients of variation of quantitative signal values (5−15%) when used to detect 8,000 to 12,000 genes. When comparing variation within and between test sites, the one-color assays demonstrated 80−95% agreement.

Although it is difficult to compare expression values generated on different platforms because of differences in labeling methods and probe sequences, MAQC was able to show good agreement between the Affymetrix® platform and the other platforms. This was particularly true when using the same biological sample (and, thus, removing variability introduced by the sample or sample preparation method). It is worth noting that Affymetrix® displayed high correlation values with RT-PCR based on comparisons of ~500 genes. The results of the MAQC report suggest that the comparisons made in this case study are valid due to the reliability of the data. Additionally, since seven out of the nine experiments in the case study were performed in the same laboratory, interlaboratory variability is not an issue with these studies. In the assessment

of consistency of findings described next, a potential source of incongruence is the decreased sensitivity for low-expression genes in the microarray platforms as compared to the gene-expression technologies and differences in probe location.

5.3. CONSISTENCY OF FINDINGS

5.3.1. Microarray Studies

An evaluation of the consistency across the four microarray studies of the testis was performed. Bowman et al. (2005) is not included because the microarray study results were not reported. In order to enhance comparability, the data from the whole testis microarray study of Plummer et al. (2007) are included in the evaluation, but the data from the microdissected regions of the testis are excluded because the lack comparison to any other study.

Eight of the nine toxicogenomic studies used the same strain, SD, and all nine used the same species (rat). Plummer et al. (2007) was the only study to use the Wistar rat strain because it is considered more susceptible to effects on the testis than SD. Table A-1 in the Appendix A includes those genes whose expression was reported to be significantly altered, as reported by Shultz et al. (2001), Thompson et al. (2005), Plummer et al. (2007) (for the whole testis only), or Liu et al. (2005) studies. Also presented in Table A-1 are the official gene names, exposure times, and directional response changes. It should be noted that some differences are to be expected in these comparisons because no two studies had identical study designs or platform, or applied the same statistical cut-offs. For example, Thompson et al. (2005) used a very short duration of exposure, whereas the other three studies had longer exposure durations. In addition, the Affymetrix®microarray platform was used only by Thompson et al. (2005) and Liu et al. (2005).

The three testis microarray studies (Thompson et al. [2005], Plummer et al. [2007], and Liu et al. [2005]) that used the "second generation chips" containing a much larger number of probes (therefore, covering many more genes) than the Clontech platform were compared. The Venn diagram, developed for these three studies, shows some unique gene expression changes for each study as well as a number of common gene expression changes (Figure 5-1). Nevertheless, significant corroboration in the direction of effect among the common genes were observed in three studies (but not with the Shultz et al. [2001, Appendix A]). Additionally, most of the genes in common were downregulated after in utero DBP exposure. Further, two genes in

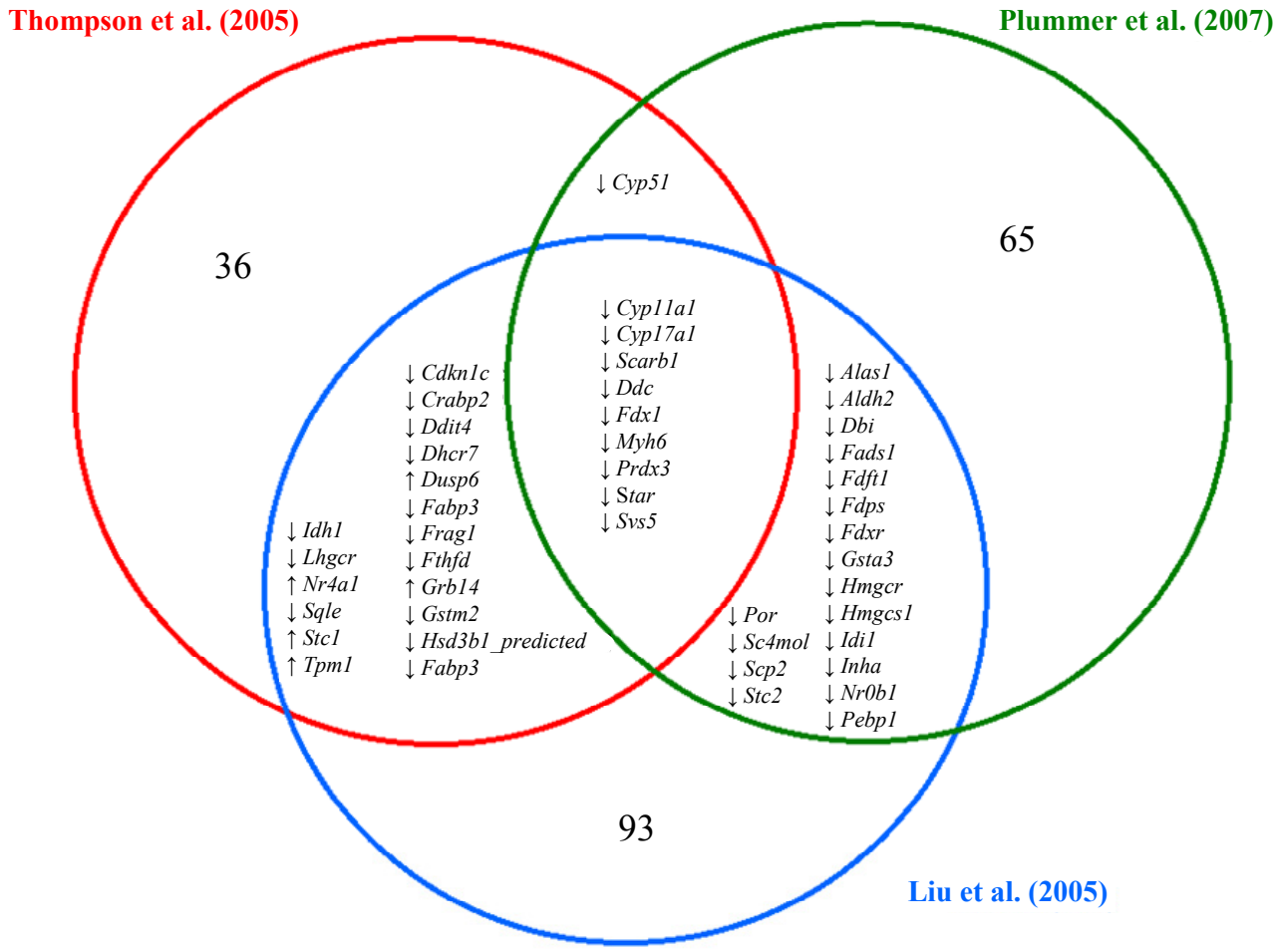

Figure 5-1. Venn diagram illustrating similarities and differences in significant gene expression changes for three of the microarray studies in the testes for three recent microarray studies: Thompson et al. (2005), Plummer et al. (2007), and Liu et al. (2005). Numbers within each circle indicate genes whose expression wasstatistically significantly altered and that are unique to the study (i.e., not replicated by either of the other two studies). Gene numbers do not include expressed sequence tags (ESTs). The red circle indicates the Thompson et al. (2005) study; the green circle indicates the Plummer et al. (2007) study; and the blue circle indicates the Liu et al. (2005) study; Black arrows indicate the direction of effect, which was the same for all three of these studies.

the steroidogenesis pathway, *Cyp11a1*, and *Scarb1*, are in common between all four microarray studies. These findings indicate that the microarray data set for DBP is relatively consistent and findings are reproducible.

A number of genes involved in steroidogenesis were found to be downregulated by DBP in all three studies (Figure 5-1). These include *Cyp 11a1, Scarb1, Star,* and *Cyp 17a1*. Other genes significantly altered include a downregulation of the serotonin and catecholamine pathway enzyme *Ddc*, the myosin, heavy polypeptide 6, cardiac muscle, alpha (*Myh6*), the androgen-regulated structural protein *Svs5*, and the cellular retinoic acid-binding protein 2 (*Crabp2*).

Other genes were significantly altered in two of the three studies. For example, in comparing the results of the two studies that utilized the same platform (Affymetrix®), the Liu et al. (2005) and Thompson et al. (2005) studies observed a downregulation of the steroidogenesis genes *Sqle* and *Hsd3b1_predicted,* cyclin-dependent protein kinase inhibitor *(Cdkn1c),* the cellular retinoic acid binding protein 2 *(Crabp2*), the FGF receptor activating protein 1 *(Frag1),* and fatty acid binding protein *(Fabp3)*. These same two studies found upregulation of the steroidogenesis gene *Nr4a3*.

There are a number of genes for which the different studies found a similar significant alteration but the direction of effect varied. For example, GSH S-transferase, mu 2 (*Gstm2*), a gene involved in xenobiotic metabolism, was found to be significantly downregulated by Liu et al. (2005) and Thompson et al. (2005) and significantly upregulated by Shultz et al. (2001). The microsomal GSH S-transferase 1 gene (*Mgst1*) was downregulated in Liu et al. (2005) and upregulated in Shultz et al. (2001). Appendix A presents a table of the significantly significant gene expression changes in the Thompson et al. (2005), Shultz et al. (2001), Liu et al. (2005), and Plummer et al. 2007 studies. These differences in microarray results can be explained by a number of factors including study differences (e.g., duration of exposure, platform, rat strain) and/or variability of microarray study results.

Overall, the data indicate that there are some unique gene expression changes for each study as well as a number of common gene expression changes. Significant corroboration in the direction of effect among the common genes was observed in at least three studies. In addition, most of the genes in common among these three studies were downregulated after in utero DBP exposure. These findings indicate that the microarray data set for DBP is very consistent and reliable although certain uncertainties remain when comparing data from different platforms with different study design.

5.3.2. Reverse Transcription-Polymerase Chain Reaction (RT-PCR) Gene Expression Findings

Comparisons were also made of RT-PCR data (Table A-2; Appendix A). All nine studies performed RT-PCR, and in the case of the Liu et al. (2005), Shultz et al. (2001), Plummer et al. (2007), and Thompson et al. (2005), the RT-PCR was performed following identification of the genes of interest in microarray studies. A number of genes were found to be similarly up- or downregulated by in utero DBP exposure. In the steroidogenesis pathway, 5 genes (*Cyp11a1, Cyp17a1, Hsd17b3, Scarb1,* and *Star*) were found to be downregulated by more than one laboratory. Some commonalities were also observed in altered gene regulation of transcription factors. *Egr1, Nfil3,* and *Nr4a1* were shown in two different studies to be upregulated. Two studies reported similar downregulation of *Nr0b1* and *Tcf1*.

Three studies (Wilson et al. [2004], Plummer et al. [2007], and Lehmann et al. [2004]) observed reduced *Insl3* gene expression. As discussed, *Insl3* has a role in sexual differentiation and testis descent. Reduced fetal *Insl3* has been shown to produce agenesis of the gubernacular ligaments. Two other genes have been shown to have DBP-induced altered expressions as assessed by RT-PCR in two laboratories: *Clu* (upregulated) and *Kit* (downregulated).

5.3.3. Protein Study Findings

All nine studies completed either Western analysis (immunoblotting) or immunohistochemistry to characterize fetal DBP-induced changes in protein expression. Usually, protein analysis was conducted for proteins that had demonstrated changes in mRNA expression. However, up- or downregulation of genes and proteins does not always occur simultaneously, so a disparity between these two experimental results is quite common. Table 5-5 presents the protein expression data.

Four proteins in the steroidogenesis pathway were shown to be downregulated by DBP exposure. These findings align well with the gene expression data presented earlier. STAR was shown to be downregulated by Western blotting in three separate experiments, and by immunolocalization in another experiment. STAR expression was found only in LCs in both the control and DBP-treated testes, with the DBP-treated testes having decreased staining intensity (Barlow et al., 2003). Quantitatively, three experiments demonstrated reduced SCARB1 protein levels in DBP-treated in fetal testes; however, immunolocalization showed

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

Official Gene Symbol	Protein Name Reported in Paper	Exposure Interval	Dose (mg/kg-d)	Method Used	Change in Protein Expression Compared to Controls	Reference
Amh	AMH	GD 12−19	500	Immunolocalization	↑ slightly in Sertoli cells	Plummer et al., 2007
Bcl2	bcl-2	GD 12−21	500	Immunolocalization	↑ in Sertoli and Leydig cells	Shultz et al., 2001
Bzrp	PBR	GD 12−19	500	Immunolocalization	↓ in interstitial cells	Lehmann et al., 2004
Cebpb	CEBPB	GD 12−19	500	Immunolocalization	↓ in interstitial cells	Liu et al., 2005
Crabp2	CRABP2 PEBP	GD 12−19	500	Immunolocalization	↓ in Leydig cells	Plummer et al., 2007
Clu	TRPM-2	GD 12−19 GD 12−21	500	Immunolocalization	↑ in Sertoli and Leydig cells	Shultz et al., 2001
Clu	TRPM-2	GD 12−19	500	Immunolocalization	↑ in Sertoli cells	Barlow et al., 2003
Cyp11a1	CYP11a1	GD 18 for 18 hrs	500	Western analysis	↓ (0.6 of control)	Thompson et al., 2005
Cyp11a1	P450ssc	GD 12−19	500	Western analysis	↓ (0.5 of control)	Lehmann et al., 2004
Cyp11a1	P450ssc	GD 12−17 or 18	500	Western analysis	↓ (0.15 at 24 hr; 0.5 at 48 hr)	Thompson et al., 2004
Cyp17a1	CYP17a1	GD 18 for 18 hrs	500	Western analysis	↓ (0.6 of control)	Thompson et al., 2005
Cyp17a1	CYP17	GD 12−17 or 18	500	Western analysis	↓ (ND at 24 hr; 0.4 of control at 48 hr)	Thompson et al., 2004
Cyp17a1	cyp17	GD 12−19	500	Western analysis	↓ (0.2 of control)	Lehmann et al., 2004
Ddc	Dopa decarboxylase	GD 12−19	500	Immunolocalization	↓ in interstitial cells	Liu et al., 2005
Grb14	GRB14	GD 12−19	500	Immunolocalization	↓ in interstitial cells and ↑ in Sertoli cells	Liu et al., 2005
Inha	INHA	GD 12−19	500	Immunolocalization	↓ in Leydig cells	Plummer et al., 2007
Insl3	Insl3	GD 12−19	500	Immunolocalization	↓ in interstitial cells	Lehmann et al., 2004
Kit	c-kit	GD 12−19	500	Immunolocalization	↓ in gonocytes	Barlow et al., 2003
Kitl	SCF	GD 12−19	500	Immunolocalization	↑ in Sertoli cells	Barlow et al., 2003

Table 5-5. (continued)

Official Gene Symbol	Protein Name Reported in Paper	Exposure Interval	Dose (mg/kg-d)	Method Used	Change in Protein Expression Compared to Controls	Reference
Nr0b1	DAX-1	GD 12−19	500	Immunolocalization	↓ in gonocytes	Liu et al., 2005
Pebp	PEBP	GD 12−19	500	Immunolocalization	↓ in Leydig cells	Plummer et al., 2007
Scarb1	SCARB1	GD 19 for 6 hrs or GD 18 for 18 hrs	500	Western analysis	↓ (0.3 of control)	Thompson et al., 2005
Scarb1	SR-B1	GD 12−17 or 18	500	Western analysis	↓ (0.15 at 24 hr; (0.7 of control at 48 hr)	Thompson et al., 2004
Scarb1	SR-B1	GD 12−19	50, 100, 500	Western analysis	↓ (0.6, 0.5, and 0.1 of control)	Lehmann et al., 2004
Scarb1	SRB1	GD 12−19	500	Immunolocalization	↓ in Leydig; ↑ in Sertoli cells	Barlow et al., 2003
Star	STAR	GD 18 for 18 hrs	500	Western analysis	↓ (0.4 of control)	Thompson et al., 2005
Star	StAR	GD 12−17 or 18	500	Western analysis	↓ (ND at 24 hr; 0.4 of control at 48 hr)	Thompson et al., 2004
Star	StAR	GD 12−19	50, 100, 500	Western analysis	↓ (0.1, 0.2, 0.1 of control)	Lehmann et al., 2004
Star	StAR	GD 12−19	500	Immunolocalization	↓ in Leydig cells	Barlow et al., 2003
Testin	testin	GD 12−19	500	Immunolocalization	↑ in Sertoli cells and gonocytes	Liu et al., 2005

ND, not detected

DBP-induced increased staining of Sertoli cells and decreased staining of Leydig cells. Both CYP11a1 and CYP17a1 protein levels were shown in several separate experiments to be reduced following DBP exposure, which correlated with microarray and PCR findings. Immunolocalization was completed for CYP11a1 and found to be downregulated in Leydig cells (Plummer et al., 2007). Using immunolocalization, CLU was found to be increased in Sertoli cells and Leydig cells. One study has shown that DBP lowers INSL3 protein immunoexpression levels in the fetal testis (McKinnell et al., 2005). The expression of SF1 was unchanged in Wistar rats, however, four proteins (CYP11a1, INHA, CRABPS, and PEBP) regulated by SF1 and AMH, were reduced in LCs following DBP exposure (Plummer et al., 2007).

5.3.4. DBP Toxicogenomic Data Set Evaluation: Consistency of Findings Summary

A comprehensive summary of the DBP toxicogenomic data set assessed in this case study, including all microarray, RT-PCR, and protein data from the nine studies, is presented in Figure 5-2. The genes and protein included in the figure are limited to those for which two or more studies detected statistically significant results. In many cases, when comparing across RT-PCR and microarray studies, a differentially expressed gene (DEG) is found in one or even several studies that is not identified in another study. For example, *Kit* was found to be downregulated in the Barlow et al. (2002), Lehmann et al. (2004), and Schultz et al. (2001) studies; by contrast, it was not altered significantly in the Liu et al. (2005) study even though it is represented on the Affymetrix® array.

Data from the Bowman et al. (2005) paper were not included because it evaluated changes in DBP-induced gene expression in the WD rather than testes. There are no other WD studies for comparisons. If an increase or decrease was reported at any time point, it was included. Multiple time points from one paper were not included, i.e., for the Thompson et al., 2005 paper studying duration of exposure, if several time points showed a change, only one was recorded as a study showing a change. For protein data, descriptions of immunohistochemical studies suggesting an increase, though without real quantitation, were still counted. For the dose-response study (Lehmann et al., 2004), data from only the 500 mg/kg-d dosing were used to allow better comparisons with the other studies.

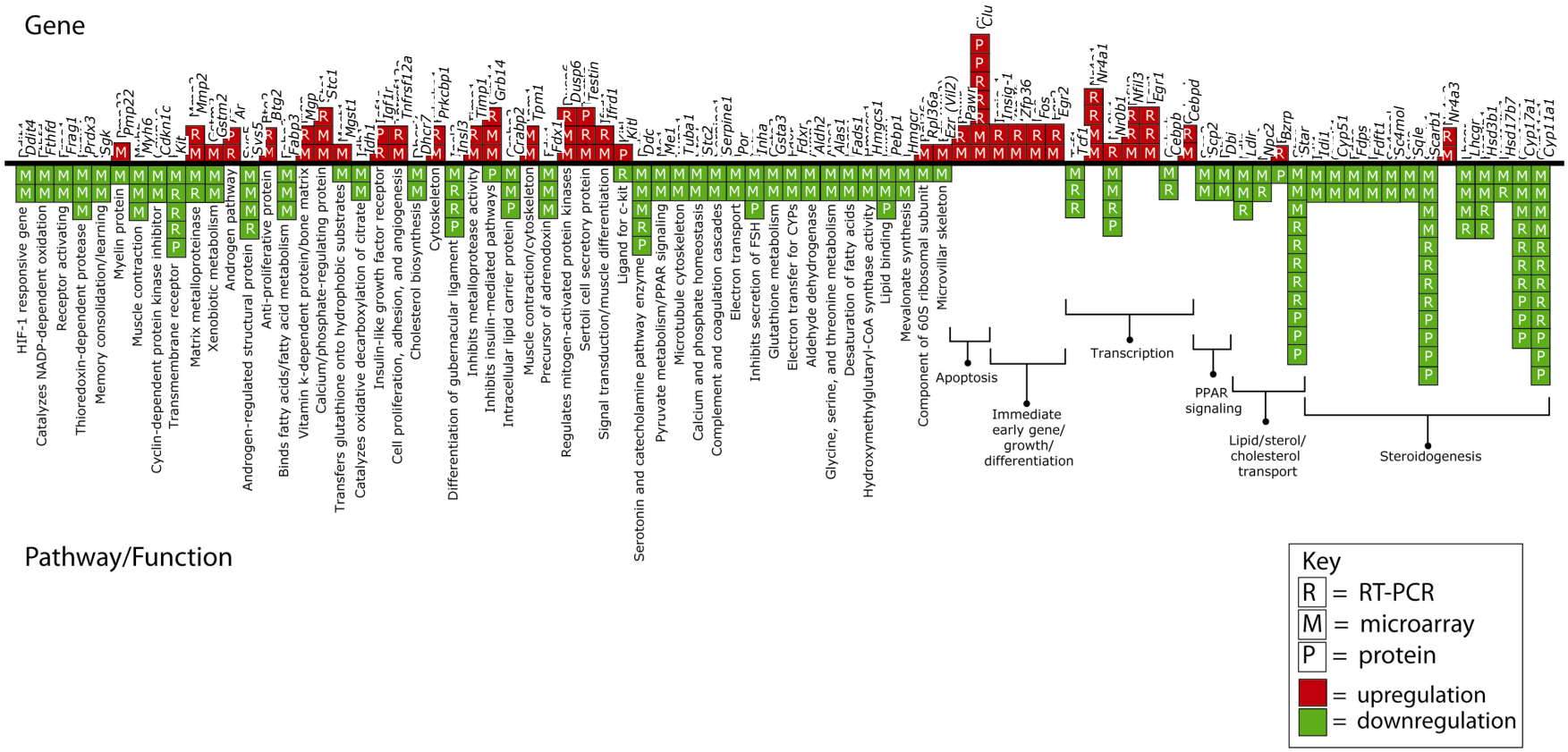

Figure 5-2. Summary of DBP-induced changes in fetal gene and protein expression. M = microarray; R = RT-PCR; P = protein. Red indicates upregulation; green indicates downregulation. Genes and protein included in the figure are limited to genes that were statistically significantly altered in two or more studies. Gene symbols are indicated at the top of the figure. The pathway or function of each gene is listed on the bottom of the figure. This information has been taken from the case study articles or from the DAVID (Database for Annotation, Visualization and Integrated Discovery http://david.abcc.ncifcrf.gov/list.jsp) entry for that gene.

Figure 5-2 presents a summary of the changes in gene and protein expression following in utero DBP exposure across the studies. What is most striking is the consistency of evidence for the DBP-induced downregulation of the steroidogenesis pathway. Both microarray and RT-PCR analysis show consistent downregulation of *Cyplla1, Cyp17a1, Star*, and *Scarb1* mRNA expression. Protein expression of *Cyplla1, Cyp17a1, Star*, and *Scarb1* is concurrently downregulated. Downregulation of *Hsd3b* and *Lhcgr* mRNA are also consistently demonstrated. Significantly, two genes involved in lipid/sterol/cholesterol transport also show downregulation: Npc2 and Ldlr. Three transcription factors (*Nfil3, Egr1,* and *Nr4a1*) demonstrate DBP-induced upregulation, while two genes (*Nr0b1* and *Tcf1*) show downregulation in a number of experiments. Three immediate early genes (*Fos, Egr2*, and *Zfp36*) are upregulated by DBP exposure. Interestingly, *Clu*, also known as T repressed prostate message-2, is upregulated, as shown by two microarray, two RT-PCR, and two protein assays.

5.4. DATA GAPS AND RESEARCH NEEDS

Based on the evaluation of nine toxicogenomic studies, a number of research needs became apparent. There are genomic data gaps for many environmental chemicals. For DBP, confirmatory RT-PCR studies for all of the genes identified from microarray studies, would give additional credence to the microarray results. Similarly, additional protein analysis, with quantitation by Western blotting and with immunolocalization, could further characterize DBP-induced effects on the male reproductive system. Looking at DBP-induced changes in gene expression in additional reproductive and nonreproductive (Thompson et al., 2005) tissues could also add information about mechanism(s) of action and tissue specificity. As testes are comprised of a number of cell types, evaluating additional homogeneous cell populations within the testes, as Plummer et al. (2007) reported, will be useful.

In order to fully consider the Case Study Question 2 (see Chapters 1 and 3), using the toxicogenomic data to determine whether there are other MOAs responsible for some of the male reproductive developmental effects, we decided that it would be helpful to analyze the raw data to assess all affected pathways. The published studies, while all excellent quality, focused their pathway analyses and descriptions on particular pathways of interest to basic science. The following chapter (Chapter 6) describes efforts to reanalyze some of the DBP microarray studies with this goal in mind.

6. NEW ANALYSES OF DBP GENOMIC STUDIES AND EXPLORATORY METHODS DEVELOPMENT FOR ANALYSIS OF GENOMIC DATA FOR RISK ASSESSMENT PURPOSES

6.1. OBJECTIVES AND INTRODUCTION

The overall goal of this chapter is to describe the new analyses of genomic and other molecular data that were performed to answer the two case study questions. The motivation for performing these new analyses is that the published DBP microarray studies were not performed for risk-assessment purposes, as is the case for the majority of the current toxicogenomic literature for all chemicals. And, some of the published analyses, such as the time course data of Thompson et al. (2005) and Plummer et al. (2007), have not yet been applied to risk assessment. This work to address the two case study questions inevitably led to the development of some new methods for analyzing genomic data for use in risk assessment. The four subobjectives of the new analyses and methods development projects were to

- Reanalyze DBP microarray data to address the Case Study Question 1: *Do the genomic data inform DBP additional MOAs and the mechanism of action for the male reproductive developmental effects?*

  We determined that it would be advantageous to reanalyze the raw data utilizing different analytical approaches (see Figure 3-1) because in most of the DBP microarray studies were analyzed to focused on further delineation of the mechanism of action relevant to one MOA, the reduction in fetal testicular T. In fact, it was the microarray and RT-PCR study results that identified the modulation of the steroidogenesis pathway as leading to reduced fetal testicular T, one of the DBP MOAs, and then, leading to a number of the male reproductive developmental effects. Not all pathways for the identified DEGs were discussed (or presented) in detail because the focus was on specific pathways of interest. A second DBP MOA of reduced *Insl3* gene expression has also been identified (Wilson et al., 2004; see Chapter 3) leading to testis descent defects. Therefore, a reanalysis that looks more broadly to define all pathways affected by DBP may inform whether there are additional pathways related to MOAs that could be linked to the unexplained male reproductive developmental outcomes caused by DBP identified in Chapter 4. Thus, the purpose of this reanalysis of the existing data set was to identify and characterize additional molecular pathways affected by DBP, beyond a reduction in fetal T and *Insl3* gene expression.

- *Explore the development of new methods for pathway analysis of microarray data for application to risk assessment.*

  The motivation was to develop methods for performing gene expression analysis of microarray data for use in risk assessment. Microarray studies for basic research purposes do not require as high a level of stringency as for risk-assessment purposes because the results are usually used for hypothesis testing (e.g., for developing an MOA hypothesis) and further studies in basic research.

- *Utilize existing DBP genomic data to develop a genetic regulatory network model, and methods for modeling, for use in risk assessment.*

  We asked whether there are data to develop a genetic regulatory network model to represent the biological interactions that are functional at different times following exposure to DBP. Regulatory network models encompass identified cellular signaling pathways from input data and, in addition, bring in gene elements that are inferred from the input data but not necessarily presented in the data. This exercise provides a more biologically enriched view of the cellular interactions inherent in the data.

- *Utilize genomic and other molecular data to address the Case Study Question 2: Do the genomic and other molecular data inform interspecies differences in MOA?*

  We utilized the available DNA, sequence, and protein similarity data to assess the rat-to-human conservation of the predicted amino acid sequences of genes involved in the steroidogenesis pathway.

The work to address the objectives of this chapter is the result of a collaborative effort between scientists at the National Center for Environmental Research STAR Center at Rutgers University and the Robert Wood Johnson Medical School UMDNJ Informatics Institute and the U.S. EPA. The analyses were performed either at Rutgers University or NHEERL-U.S. EPA.

The work presented in this chapter is highly technical and thus, is intended to be beneficial to scientists with expertise in genomic and genetic data analysis. The technical details of the analyses are provided in order that scientists could apply these methods to their work. Such an approach will allow the risk assessor proficient in microarray analysis methodology an opportunity to apply these methods. The last section of this chapter (section 6.6.) summarizes the findings for a scientific audience without a strong understanding of microarray analysis methods..

6.2. REANALYSIS OF GENE EXPRESSION DATA TO IDENTIFY NEW MOAs TO ELUCIDATE UNEXPLAINED TESTICULAR DEVELOPMENT ENDPOINTS AFTER IN UTERO DBP EXPOSURE

6.2.1. Objective of the Reanalysis of the Liu et al. (2005) Study

The goal was to reanalyze DBP microarray data to address the Case Study Question 1: *Do the genomic data inform DBP additional MOAs and the mechanism of action for the male reproductive developmental effects?* Modulation of the steroidogenesis pathway, leading to reduced fetal testicular T, has been identified from the microarray and RT-PCR studies as one MOA for DBP's male reproductive developmental effects. The Liu et al. (2005) study focused on the steroidogenesis and related pathway. Not all pathways for the identified DEGs were discussed (or presented) in detail because the focus of the study was on steroidogenesis. Therefore, a reanalysis that looks more broadly to define all pathways affected by DBP may inform whether there are additional modes and mechanisms of action that could be linked to the unexplained male reproductive developmental outcomes caused by DBP identified in Chapter 4. The purpose for the reanalysis of the existing data sets is to identify and characterize additional molecular pathways affected by DBP, beyond the effects on the androgen-mediated male reproductive developmental toxicity pathways.

The Liu et al. (2005) study was selected for reanalysis because the data set had a comprehensive exposure scenario that covered the critical window for developmental exposure to DBP (GD 12−19), the Affymetrix® chip was used (compatible with the proprietary and free software programs used for pathway level analysis), and the data were provided by Dr. Kevin Gaido, a collaborator on this project. Some limitations of the Liu et al. (2005) data set are the small number of samples (i.e., 3 controls and 3 DBP-treated) and the within sample variance. This study was a comparative analysis of six phthalate esters. However, only the DBP treatment and vehicle control data were used for this analysis.

The Liu et al. (2005) study investigated global gene expression in the fetal testis following in utero exposure to a series of phthalate esters, including both developmentally toxic phthalates (DBP, BBP, DPP, and DEHP) and non-developmentally toxic phthalates (DMP, DEP, and DOTP) (Liu et al., 2005). The original analysis was based on a two-way nested ANOVA model using Bonferroni correction that identified 391 significantly expressed genes from the

control out of the approximately 30,000 genes queried. In their analysis, two classes of phthalate esters were distinguished based on the gene expression profiles. The authors also showed that developmentally toxic phthalates targeted gene pathways associated with steroidogenesis, lipid and cholesterol homeostasis, insulin signaling, transcriptional regulation, and oxidative stress.

The common approach of interrogating a handful of genes that pass user-defined statistical filtering criterion to understand the biology of a system has some limitations (Tomfohr et al., 2005). These include the following:

- Often times after correcting for multiple hypothesis testing, few or no genes pass the threshold of statistical significance because the biological variances are modest relative to the noise inherent in a microarray experiment.

- Alternatively, one is left with a long list of statistically significant genes that have no unifying biological theme. Interpretation of such a list can be daunting.

- Additionally, since cellular processes are not affected by changes in single genes, but a set of genes acting in concert, single gene analysis can miss out on relevant biological information.

- Often times, there is little concordance between lists of statistically significant genes from similar studies conducted by two groups.

6.2.1.1. *Differentially Expressed Gene (DEG) Identification: Linear Weighted Normalization*

The data set for the vehicle treated and DBP treated samples were input into the proprietary software Rosetta Resolver®. A principal component analysis (PCA) of the entire data set shows a distinct treatment response (i.e., the control and treated samples separate out clearly into two distinct groups [see Figure 6-1]). Additionally, it demonstrates certain limitations of this data set—namely the variance in the data set between similarly treated samples. This is apparent from the fact that even though the two groups show separation along two different axes, they are not tightly grouped together in space.

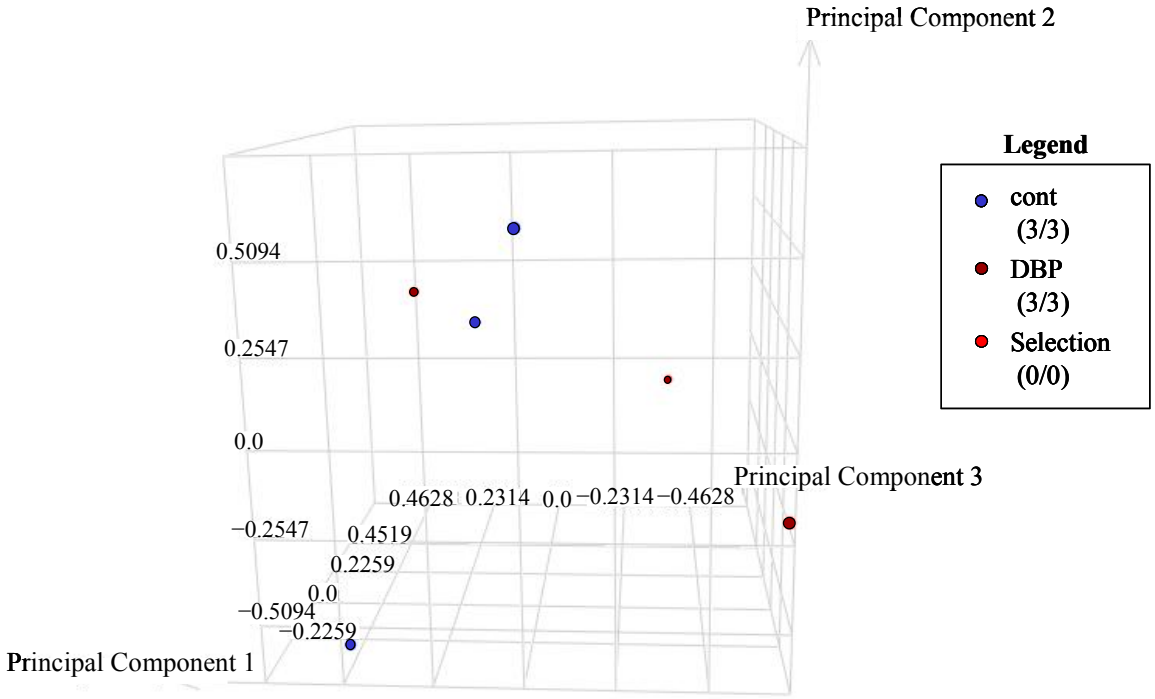

Figure 6-1. Principal component analysis (PCA) representation of Liu et al. (2005) data set. PCA is a standard technique for visualization of complex data, showing the distribution of each sample and the degree of similarity to one another. PCA shows relationship of all six samples, DBP-treated (red) and concurrent vehicle control (blue). Generated by Rosetta Resolver Software v 7.0.

Next, the gene expression data were normalized using a linear weighted model in Rosetta Resolver®. The Rosetta Resolver® system is a comprehensive gene expression analysis solution that incorporates powerful analysis tools with a robust, scalable database. The annotated genes of the rat genome on the Affymetrix® gene chip, ~30,000 genes, were input into the significance analysis with a Benjamini Hochberg Multiple FDR correction for multiple testing applied at $p < 0.01$, a more stringent statistical cut-off. Of the ~30,000 genes, the analysis passed 118 genes as being significantly altered following DBP exposure. Of these, 17,496 genes did not pass the statistical filter, and 13,428 genes were not affected by the treatment. One possible reason that only 118 genes passed the multiple-testing correction filter is that there is a high variance between individual samples as demonstrated by the PCA.

Using the linear-weighted normalization analysis, we relaxed the filtering criterion to include more genes because the objective of this exercise was to identify additional pathways affected by DBP, and starting out with 118 genes would be limiting in that regard. Additionally, often times, researchers have to make a judgment call on when to put emphasis on statistical significance and when to focus on the biological significance. Since the objective was to use the

gene expression data to gain new information about DBP toxicity, we deemed it suitable to relax the statistical filtering criteria to obtain maximum numbers of genes to upload to pathway mapping software.

The next filtering strategy involved applying a statistical t-test of $p < 0.05$ (see Figure 6-2) and did no multiple testing correction (MTC) was applied because when MTC was applied, no genes were identified as significant. Of the 31,000 gene probes present on the RAE A and B Affymetrix® GeneChips®, 1,977 passed this filter.

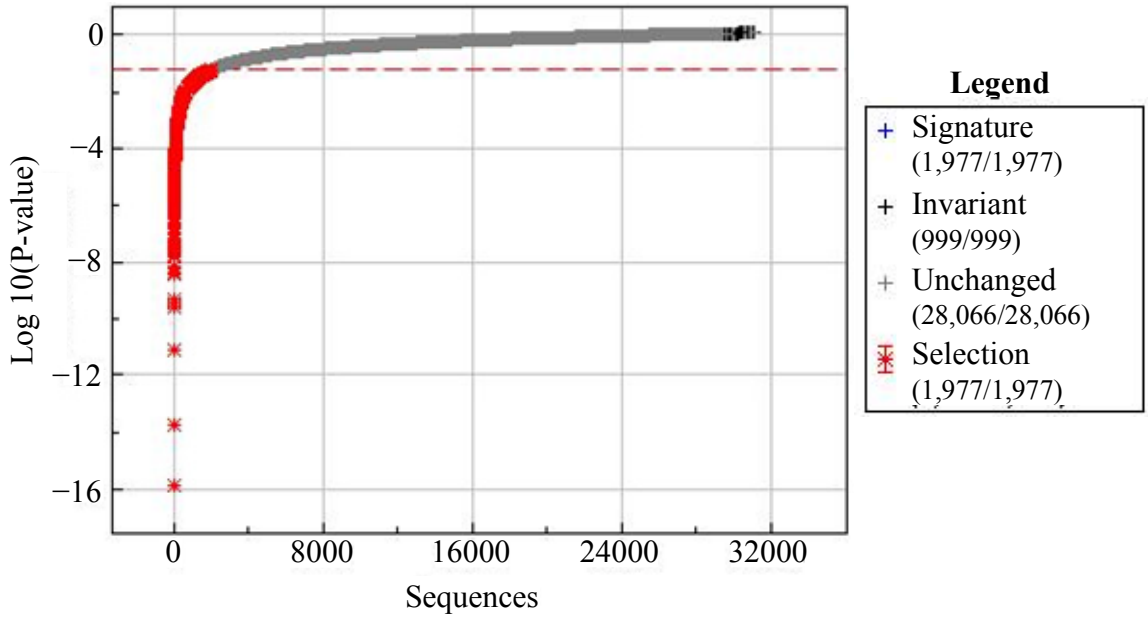

Figure 6-2. Selection of significant genes using Rosetta Resolver®. 30,000 genes were input into the significance analysis with a *p*-value cutoff of 0.05. No multiple-testing correction was applied. The analysis passed 1,977 genes as being significantly altered following DBP exposure. Of the remaining genes, 999 genes were differentially expressed but did not pass the statistical filter, and 28,066 genes were not affected by the treatment.

The set of 1,977 genes was deemed suitable to perform a comprehensive pathway-level analysis because about one third of the DEGs (999) did not meet the statistical cut-off criteria (a p value cutoff < or = 0.05). To do this, the list of 1,977 genes was inputted into a second software program called GeneGo. GeneGo is a leading provider of data analysis solutions in systems biology. Its proprietary database MetaCore^{TM}'s sophisticated analytical tools enable the identification and prioritization of the most relevant pathways, networks, and cellular processes affected by a given condition.

6.2.1.2. *Differentially Expressed Gene (DEG) Identification: Signal-to-Noise Ratio (SNR)*

We also identified DEGs two independent methods, Signal-to-Noise Ratio (SNR) (Golub et al., 1999) and a two sample t-test from the Liu et al. (2005 DBP data. SNR reflects the difference between the classes relative to the standard deviation within the classes.

Equation 6-1 evaluates the means and standard deviations of the expression levels of gene *g* for the samples in group 1 (vehicle control) and group 2 (DBP treated), respectively.

For a given gene (*g*) we evaluate the SNR using Equation 6-1:

$$SNR = \frac{\left|\mu_1(g) - \mu_2(g)\right|}{\sigma_1(g) + \sigma_2(g)} \qquad (6\text{-}1)$$

The means and standard deviations of the expression levels of gene *g* are denoted with μ and σ, respectively, for the samples in group 1 (vehicle control) and group 2 (DBP treated).

A high value of SNR is indicative of a strong distinction between the groups—i.e., vehicle and DBP treated. In order to identify the DEGs whose expression was altered by DBP, 1,000 random gene expressions were permutated from the whole data set, and their SNR was computed. The ratio of the randomly generated SNR value that is higher than the actual SNR value determined whether the expression of the probe set is differentially expressed or not. Appendix B lists the algorithm for selecting DEGs (see Figure B-1). A list of 1,559 probe sets was identified as being differentially expressed following a statistical cut-off of $p < 0.05$. The heat map (see Figure 6-3) illustrates the distinction between the vehicle and treated samples. On the other hand, Student's t-test ($p < 0.05$) revealed 1,876 probe sets being statistically significant.

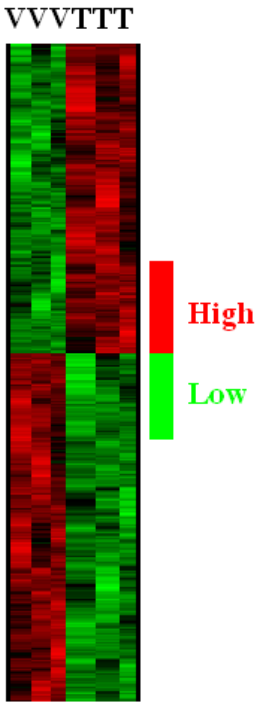

Figure 6-3. Heat map of 1,577 differentially expressed genes (DEGs) from SNR analysis method. V = Vehicle, T = Treated samples. Data used for analysis from Liu et al. (2005). Columns represent the six treatment conditions (3 DBP treatments, 3 vehicle controls). Rows represent the different 1,577 DEGs. Red represents up regulation of gene expression, and green represents down regulation of gene expression.

ArrayTrack was used to calculate the pathway enrichment for the two DEGs lists, SNR list, and t-test list. To investigate interactions of genes at the pathway level, the Kyoto Encyclopedia of Genes and Genomes (KEGG; http://www.genome.jp/kegg) database was utilized as a pathway mapping tool, and a Fisher's exact test was used to compute the significance. The top five enriched pathways as derived from both gene lists are common: biosynthesis of steroids, terpenoid biosynthesis, GSH biosynthesis, and carbon fixation. The SNR gene list maps to more pathways than the t-test gene list, even though the number of DEGs was greater in the t-test generated gene list than in the SNR gene list.

6.2.2. Pathway Analysis of Liu et al. (2005) Comparing Two Methods

Pathway-analysis methods and software have been previously developed for analysis of microarray data for basic and applied research. Pathway-level analysis mainly depends on the definition of the pathways (database) and significance level uses to measure the differential expressions. Using these validated methods, a pathway analysis was performed. Further, a comparison of methods between the results from using different analytical approaches, SNR and linear weighted normalization, was performed.

Analysis of DBP toxicogenomic studies was carried out using many proprietary databases and software packages that are available to the microarray community with enhanced bioinformatic capabilities for pathway and functional level analysis (Rosetta Resolver®, MetaCore™ GeneGo, Ingenuity® Pathway Knowledgebase. These software tools accept lists of genes of interest and then using their database of knowledge about these gene elements, map them to cellular pathways known to exist from experimental literature. The advantage of trying to understand groups of genes acting in a similar cellular process such as cell cycle provides more meaningful results as opposed to trying to understand one gene at a time, which may have no relationship to other genes on a statistically filtered list. The rationale behind the exercise was that interrogation of multiple databases would result in a more complete mining of the microarray data sets, which may provide an understanding of all of the potential DBP MOAs underlying the testes reproductive developmental effects. Analysis using different statistical tools provides information about the similarities and differences in results.

Figure 6-4 shows the schematic of the comparative analysis protocol. The GeneGo analysis normalized data set revealed that 131 biological processes ($p < 0.05$) were associated with the 1,977 DEGs. Table 6-1 lists the pathways with a $p < 0.05$ (Fisher exact t-test). Comparisons made on the level of gene lists obtained by different statistical methods often do not converge (Stocco et al., 2005). We decided to perform a comparison of methods based on the assumption that biologically related groups of genes, such as metabolic or signaling pathways, may be more valid if identified using different microarray analysis methods. Towards this effort, we treated the gene list (1,559 genes) using SNR to a pathway level analysis using GeneGo, similar to the analysis performed on the linear weighted normalization results. Table 6-2 lists the result of this analysis.

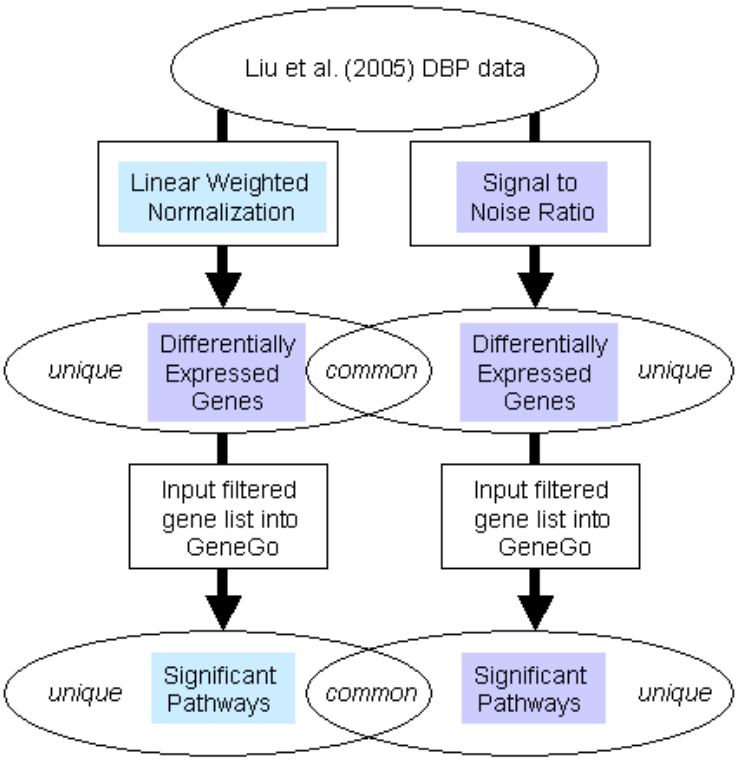

Figure 6-4. Schematic of the two analysis methods (linear weighted normalization and SNR) for identifying differentially expressed genes and subsequent pathway analysis using GeneGo. Two separate analyses, linear weighted normalization and SNR statistical filters, were performed to identify common and unique genes from the Liu et al. (2005) data. The two separate filtered gene lists were input into GeneGo to identify statistically significantly affected pathways. Common and unique pathway lists were generated.

Table 6-1. GeneGo pathway analysis of significant genes affected by DBP

Pathway	Biological process	p-Value[a]	No. of genes[b,c]
NF-AT signaling in cardiac hypertrophy	Disease	2.23E−04	19/90
MIF—the neuroendocrine-macrophage connector	Immune response	3.00E−04	19/92
Lysine metabolism	Amino acid metabolism	3.05E−04	9/27
Cholesterol metabolism	Steroid metabolism	6.95E−04	6/14
Glycolysis and gluconeogenesis (short map)	Carbohydrates metabolism	7.40E−04	10/36
Integrin-mediated cell adhesion	Cell adhesion	8.44E−04	18/92
Tryptophan metabolism	Amino acid metabolism	9.56E−04	9/31
Cholesterol biosynthesis	Steroid metabolism	1.44E−03	7/21
ECM remodeling	Cell adhesion	1.64E−03	13/60
Regulation of lipid metabolism via PPAR, RXR, and VDR	Transcription	1.96E−03	7/22
Propionate metabolism p.2	Carbohydrates metabolism	1.96E−03	7/22
PPAR regulation of lipid metabolism	Regulation of lipid metabolism	2.04E−03	8/28
Mitochondrial long chain fatty acid beta-oxidation	Lipid metabolism	2.28E−03	6/17
Role of VDR in regulation of genes involved in osteoporosis	Transcription	3.16E−03	12/57
ChREBP regulation pathway	G-protein coupled receptor signaling	3.82E−03	10/44
Androstenedione and testosterone biosynthesis and metabolism p.1	Steroid metabolism	4.30E−03	6/19
Arginine metabolism	Amino acid metabolism	4.45E−03	9/38
Regulation of fatty acid synthesis: NLTP and EHHADH	Regulation of lipid metabolism	5.02E−03	4/9
Angiotensin signaling via STATs	Growth and differentiation	5.18E−03	11/53
Cytoskeleton remodeling	Cell adhesion	5.19E−03	26/176
dGTP metabolism	Nucleotide metabolism	5.34E−03	9/39
TCA	Amino acid metabolism	5.70E−03	6/20
Glycolysis and gluconeogenesis p. 1	Carbohydrates metabolism	5.70E−03	6/20
Peroxisomal branched chain fatty acid oxidation	Lipid metabolism	5.70E−03	6/20

1

Table 6-1. (continued)

Pathway	Biological process	p-value[a]	No. of genes[bc]
Gamma-aminobutyrate (GABA) biosynthesis and metabolism	Metabolism of mediators	5.70E−03	6/20
Ligand-dependent activation of the ESR1/SP pathway	Response to hormone stimulus	6.38E−03	9/40
Integrin inside-out signaling	Cell adhesion	6.85E−03	14/78
Reverse signaling by ephrin B	Cell adhesion	6.86E−03	15/86
G-protein beta/gamma signaling cascades	G-protein coupled receptor protein signaling pathway	6.94E−03	11/55
Activation of PKC via G-Protein coupled receptor	G-protein coupled receptor protein signaling pathway	7.65E−03	15/87
Gap junctions	Cell adhesion	8.51E−03	10/49
WNT signaling pathway	Proteolysis	8.59E−03	7/28
Angiotensin activation of ERK	G-protein coupled receptor protein signaling pathway	9.12E−03	11/57
Role of Akt in hypoxia induced HIF1 activation	Proteolysis	9.83E−03	10/50
Regulation of actin cytoskeleton by Rho GTPases	Small GTPase mediated signal transduction	1.18E−02	11/59
CCR3 signaling in eosinophils	Immune response	1.22E−02	18/117
MAG-dependent inhibition of neurite outgrowth	Response to extracellular stimulus	1.47E−02	10/53
Endothelial cell contacts by junctional mechanisms	Cell adhesion	1.80E−02	7/32
Fructose metabolism	Carbohydrates metabolism	1.80E−02	7/32
Regulation of lipid metabolism via LXR, NF-Y and SREBP	Transcription	1.80E−02	7/32
CXCR4 signaling pathway	Cytokine and chemokine mediated signaling pathway	1.89E−02	10/55
Serotonin-melatonin biosynthesis and metabolism	Metabolism of mediators	2.04E−02	5/19
Glycolysis and gluconeogenesis p. 2	Carbohydrates metabolism	2.15E−02	4/13
Oxidative phosphorylation	Energy metabolism	2.37E−02	15/99
Urea cycle	Amino acid metabolism	2.58E−02	6/27

Table 6-1. (continued)

Pathway	Biological process	p-value[a]	No. of genes[bc]
G-proteins mediated regulation p.38 and JNK signaling	G-protein coupled receptor protein signaling pathway	2.60E−02	11/66
Transcription factor tubby signaling pathways	Transcription	2.63E−02	8/42
Role PKA in cytoskeleton reorganization	Protein kinase cascade	2.64E−02	13/83
Ephrins signaling	Cell adhesion	2.66E−02	10/58
Propionate metabolism p.1	Carbohydrates metabolism	2.81E−02	4/14
Estrone metabolism	Steroid metabolism	2.81E−02	4/14
Regulation of acetyl-CoA carboxylase 2 activity in muscle	Response to extracellular stimulus	2.81E−02	4/14
Chemokines and adhesion	Cytokine and chemokine mediated signaling pathway	2.82E−02	23/174
Arachidonic acid production	Lipid metabolism	2.87E−02	7/35
dCTP/dUTP metabolism	Nucleotide metabolism	2.99E−02	8/43
Regulation of lipid metabolism by niacin and isoprenaline	Regulation of lipid metabolism	3.01E−02	9/51
Ubiquinone metabolism	Vitamin and cofactor metabolism	3.01E−02	9/51
Phenylalanine metabolism	Amino acid metabolism	3.05E−02	6/28
Leptin signaling via JAK/STAT and MAPK cascades	Response to hormone stimulus	3.57E−02	6/29
IMP biosynthesis	Nucleotide metabolism	3.70E−02	3/9
EPO-induced Jak-STAT pathway	Response to extracellular stimulus	3.78E−02	7/37
Integrin outside-in signaling	Cell adhesion	3.95E−02	12/79
Brca1 as transcription regulator	Cell cycle	4.15E−02	6/30
P53 signaling pathway	Transcription regulation	4.28E−02	8/46
Bile acid biosynthesis	Steroid metabolism	4.43E−02	5/23
Histidine-glutamate-glutamine and proline metabolism	Amino acid metabolism	4.79E−02	8/47
NTS activation of IL-8 in colonocytes	Immune response	4.85E−02	10/64

[a]Ordered from most significant (lowest *p*-value) to less significant.
[b]Number of genes from the DBP-exposed gene list mapping to the GeneGo pathway.
[c]Total number of genes in the GeneGo pathway.

Table 6-2. Significant biological pathways corresponding to differentially expressed genes (DEGs) obtained from SNR analysis input into GeneGo

Pathway	Biological Process	p-Value[a]	No. of genes[bc]
Cholesterol Biosynthesis	Steroid metabolism	1.81E−09	15/21
Propionate metabolism p.2	Carbohydrates metabolism	5.54E−06	12/22
MIF—the neuroendocrine-macrophage connector	Immune response	3.22E−04	25/92
Tryptophan metabolism	Amino acid metabolism	3.78E−04	12/31
Lysine metabolism	Amino acid metabolism	3.93E−04	11/27
Cholesterol metabolism	Steroid metabolism	1.09E−03	7/14
NF-AT signaling in cardiac hypertrophy	Disease	1.38E−03	23/90
Glycolysis and gluconeogenesis (short map)	Carbohydrates metabolism	1.77E−03	12/36
G-alpha(q) regulation of lipid metabolism	Regulation of lipid metabolism	1.93E−03	13/41
Activation of PKC via G-protein coupled receptor	G-proteins/GPCR	2.00E−03	22/87
Fructose metabolism	Carbohydrates metabolism	2.06E−03	11/32
Regulation of lipid metabolism by niacin and isoprenaline	Regulation of lipid metabolism	2.08E−03	15/51
ATP metabolism	Nucleotide metabolism	2.09E−03	16/56
Angiotensin activation of ERK	Growth and differentiation	2.55E−03	16/57
NTS activation of IL-8 in colonocytes	Immune response	3.60E−03	17/64
Leucine, isoleucine, and valine metabolism.p.2	Amino acid metabolism	3.64E−03	9/25

Table 6-2. (continued)

Pathway	Biological Process	p-Value[a]	No. of genes[bc]
Reverse signaling by ephrin B	Cell adhesion	3.92E−03	21/86
Cortisone biosynthesis and metabolism	Steroid metabolism	4.31E−03	7/17
CXCR4 signaling pathway	Immune response	4.63E−03	15/55
G-Protein beta/gamma signaling cascades	G-proteins/GPCR	4.63E−03	15/55
Glutathione metabolism	Vitamin and cofactor metabolism	5.77E−03	11/36
Mitochondrial ketone bodies biosynthesis and metabolism	Lipid metabolism	5.96E−03	5/10
Integrin inside-out signaling	Cell adhesion	6.07E−03	19/78
Propionate metabolism p.1	Carbohydrates metabolism	6.51E−03	6/14
Role of VDR in regulation of genes involved in osteoporosis	Transcription factors	6.63E−03	15/57
Endothelial cell contacts by junctional mechanisms	Cell adhesion	7.02E−03	10/32
EPO-induced Jak-STAT pathway	Cell survival	7.24E−03	11/37
A3 receptor signaling	G-proteins/GPCR	8.08E−03	19/80
Angiotensin signaling via STATs	Growth and differentiation	8.28E−03	14/53
MAG-dependent inhibition of neurite outgrowth	Growth and differentiation	8.28E−03	14/53
Phenylalanine metabolism	Amino acid metabolism	8.48E−03	9/28
Androstenedione and testosterone biosynthesis and metabolism p.1	Steroid metabolism	8.76E−03	7/19
Cytoskeleton remodeling	Cell adhesion	9.69E−03	35/176

Table 6-2. (continued)

Pathway	Biological Process	p-Value[a]	No. of genes[bc]
ChREBP regulation pathway	Regulation of transcription	1.08E−02	12/44
Leptin signaling via JAK/STAT and MAPK cascades	Growth and differentiation	1.09E−02	9/29
dGTP metabolism	Nucleotide metabolism	1.10E−02	11/39
TCA	Amino acid metabolism	1.20E−02	7/20
Glycolysis and gluconeogenesis p. 1	Carbohydrates metabolism	1.20E−02	7/20
Gamma-aminobutyrate (GABA) biosynthesis and metabolism	Metabolism of mediators	1.20E−02	7/20
BAD phosphorylation	Apoptosis	1.21E−02	19/83
Ligand-dependent activation of the ESR1/SP pathway	Hormones	1.34E−02	11/40
RAB5A regulation pathway	G-proteins/RAS-group	1.49E−02	5/12
Integrin outside-in signaling	Cell adhesion	1.50E−02	18/79
Hedgehog and PTH signaling pathways participation in bone and cartilage development	Growth and differentiation	1.62E−02	11/41
G-Proteins mediated regulation MARK-ERK signaling	G-proteins/GPCR	1.64E−02	17/74
Integrin-mediated cell adhesion	Cell adhesion	1.78E−02	20/92
Mitochondrial long chain fatty acid beta-oxidation	Lipid metabolism	1.88E−02	6/17
CCR3 signaling in eosinophils	Immune response	2.02E−02	24/117
Regulation of lipid metabolism via PPAR, RXR, and VDR	Transcription factors	2.07E−02	7/22

Table 6-2. (continued)

Pathway	Biological Process	p-Value[a]	No. of genes[bc]
Glycolysis and gluconeogenesis p. 2	Carbohydrates metabolism	2.16E−02	5/13
Regulation of fatty acid synthesis: NLTP and EHHADH	Regulation of lipid metabolism	2.30E−02	4/9
Role PKA in cytoskeleton reorganization	Kinases	2.43E−02	18/83
Arginine metabolism	Amino acid metabolism	2.44E−02	10/38
ECM remodeling	Cell adhesion	2.45E−02	14/60
Ca ($^{2+}$)-dependent NF-AT signaling in cardiac hypertrophy	Disease	2.55E−02	15/66
WNT signaling pathway	Growth and differentiation	2.64E−02	8/28
PPAR regulation of lipid metabolism	Regulation of lipid metabolism	2.64E−02	8/28
Insulin regulation of the protein synthesis	Translation regulation	2.67E−02	13/55
CXCR4 signaling via second messenger	Immune response	2.67E−02	13/55
Angiotensin signaling via beta-Arrestin	Growth and differentiation	2.71E−02	11/44
Estrone metabolism	Steroid metabolism	2.99E−02	5/14
Regulation of acetyl-CoA carboxylase 2 activity in muscle	Growth and differentiation	2.99E−02	5/14
Prolactin receptor signaling	Growth factors	3.19E−02	14/62
Triacylglycerol metabolism p.1	Lipid metabolism	3.23E−02	8/29
Serotonin-melatonin biosynthesis and metabolism	Metabolism of mediators	3.27E−02	6/19
Angiotensin signaling via PYK2	Growth and differentiation	3.32E−02	16/74

Table 6-2. (continued)

Pathway	Biological Process	p-Value[a]	No. of genes[bc]
G-Protein alpha-i signaling cascades	G-proteins/GPCR	3.36E−02	12/51
dATP/dITP metabolism	Nucleotide metabolism	3.86E−02	12/52
Brca1 as transcription regulator	Cell-cycle control	3.90E−02	8/30
Ephrins signaling	Cell adhesion	3.99E−02	13/58
Mitochondrial unsaturated fatty acid beta-oxidation	Lipid metabolism	4.01E−02	5/15
GDNF signaling	Growth and differentiation	4.08E−02	7/25
Aspartate and asparagine metabolism	Amino acid metabolism	4.15E−02	6/20
Peroxisomal branched chain fatty acid oxidation	Lipid metabolism	4.15E−02	6/20
Histidine-glutamate-glutamine and proline metabolism	Amino acid metabolism	4.24E−02	11/47
TGF-beta receptor signaling	Growth and differentiation	4.51E−02	13/59
Regulation of actin cytoskeleton by Rho GTPases	G-proteins/RAS-group	4.51E−02	13/59
G-Protein alpha-s signaling cascades	G-proteins/GPCR	4.51E−02	13/59
A1 receptor signaling	G-proteins/GPCR	4.61E−02	16/77
Membrane-bound ESR1: interaction with growth factors signaling	Growth and differentiation	4.64E−02	10/42
Transcription factor Tubby signaling pathways	Regulation of transcription	4.64E−02	10/42
Histamine metabolism	Metabolism of mediators	4.83E−02	4/11
PPAR pathway	Transcription factors	4.86E−02	11/48

Table 6-2. (continued)

Pathway	Biological Process	p-Value[a]	No. of genes[bc]
Cross-talk VEGF and angiopoietin 1 signaling	Growth and differentiation	5.08E−02	9/37
EPO-induced MAPK pathway	Growth and differentiation	5.08E−02	13/60

[a]Ordered from most significant (lowest *p*-value) to less significant.
[b]Number of genes from the DBP exposed gene list mapping to the GeneGo pathway.
[c]Total number of genes in the GeneGo pathway.

Table 6-3 lists the pathways that are in common between conducting the two different analyses by using the GeneGo analysis (i.e., the union of the two separate pathway lists; see Tables 6-1 and 6-2). This analysis highlights biological processes and pathways that are affected by DBP exposure to fetal testis besides the already established changes in the steroidogenesis pathway. An attempt to link these unique pathways and processes to the DBP-induced male reproductive toxicity outcomes will be made based on the published literature.

Cholesterol biosynthesis/metabolism and associated pathways underlie one of the MOAs of DBP. To determine a metric for statistical analysis protocols of toxicogenomic data, we chose to compare the genes that are involved in the cholesterol biosynthesis/metabolism as identified by the three independent analysis methods (described herein) as well as the published data set from Liu et al. (2005) (see Table 6-4). These results show that there is a high degree of overlap in the most biologically relevant pathway/process involved in DBP toxicity, even when different statistical procedures are used for analysis of the same data set. These are in agreement with the published literature, giving the approaches used in this exercise biological confidence.

By utilizing databases such as GeneGo, additional canonical pathways and biological processes were identified that may play an important role in its toxicity. Regulation of steroidogenesis requires multiple signaling pathways and growth factors (Stocco et al., 2005). Signaling pathways, like the protein kinase C pathway, arachidonic acid metabolism, growth factors, chloride ion, and the calcium messenger system are capable of regulating/modulating steroid hormone biosynthesis. It is possible that some of the pathways and processes identified by the two methods may play a role in the regulation of steroidogenesis, known to be affected by DBP. Another scenario could be that these pathways and processes have yet to be associated with DBP-induced toxicity.

Previous transcriptional studies have been shown that DBP does not bind to the AR unlike flutamide (Parks et al., 2000), rather, it interrupts T synthesis (Shultz et al., 2001). The androstenedione and T biosynthesis and metabolism pathway was one of the common pathways in the GeneGo analysis of the two different methods gene lists (see Figures 6-5 and 6-6). We investigated the potential role of AR in DBP-induced toxicity by querying the GeneGo database based on the transcriptional profiling data.

Table 6-3. Common pathways between the linear weighted normalization and SNR analyses of differentially expressed genes (DEGs) after in utero DBP exposure from the Liu et al. (2005) data[a,b,c]

Biological Process	Pathways
Cell adhesion	Cytoskeleton remodeling
	ECM remodeling
	Endothelial cell contacts by junctional mechanisms
	Ephrins signaling
	Integrin inside-out signaling
	Integrin outside-in signaling
	Integrin-mediated cell adhesion
	Reverse signaling by ephrin B
Cell signaling	Activation of PKC via G-Protein coupled receptor
	CCR3 signaling in eosinophils
	ChREBP regulation pathway
	G-Protein beta/gamma signaling cascades
	G-Proteins mediated regulation p. 38 and JNK signaling
	Leptin signaling via JAK/STAT and MAPK cascades[2]
	Regulation of actin cytoskeleton by Rho GTPases
	Role PKA in cytoskeleton reorganization
Disease	NF-AT signaling in cardiac hypertrophy
	NTS activation of IL-8 in colonocytes
Growth and differentiation	Angiotensin activation of ERK
	Angiotensin signaling via STATs
	EPO-induced Jak-STAT pathway
	MAG-dependent inhibition of neurite outgrowth
	Regulation of acetyl-CoA carboxylase 2 activity in muscle
	WNT signaling pathway
Hormones	Ligand-dependent activation of the ESR1/SP pathway
Immune response	MIF - the neuroendocrine-macrophage connector
	CXCR4 signaling pathway

1

Table 6-3. (continued)

Biological Process	Pathways
Metabolism	***Androstenedione and testosterone biosynthesis and metabolism p.1[2]***
	Cholesterol biosynthesis[2]
	Cholesterol metabolism[2]
	dATP/dITP metabolism
	dGTP metabolism
	Estrone metabolism
	Fructose metabolism
	G-alpha(q) regulation of lipid metabolism
	Gamma-aminobutyrate (GABA) biosynthesis and metabolism
	Glutathione metabolism
	Glycolysis and gluconeogenesis (short map)
	Glycolysis and gluconeogenesis p. 1
	Glycolysis and gluconeogenesis p. 2
	Histamine metabolism
	Histidine-glutamate-glutamine and proline metabolism
	Leucine, isoleucine and valine metabolism p. 2
	Lysine metabolism
	Mitochondrial ketone bodies biosynthesis and metabolism
	Mitochondrial long chain fatty acid beta-oxidation
	Mitochondrial unsaturated fatty acid beta-oxidation
	Peroxisomal branched chain fatty acid oxidation
Metabolism	Phenylalanine metabolism
	PPAR regulation of lipid metabolism[2]
	Propionate metabolism p.1[2]
	Propionate metabolism p.2[2]
	Regulation of fatty acid synthesis: NLTP and EHHADH
	Regulation of lipid metabolism by niacin and isoprenaline
	Regulation of lipid metabolism via LXR, NF-Y, and SREBP[2]
	Regulation of lipid metabolism via PPAR, RXR, and VDR[2]
	Serotonin—melatonin biosynthesis and metabolism
	TCA
	Triacylglycerol metabolism p.1
	Tryptophan metabolism

Table 6-3. (continued)

Biological Process	Pathways
Transcription	Brca1 as transcription regulator
	Role of VDR in regulation of genes involved in osteoporosis
	Transcription factor Tubby signaling pathways

[a]Significant gene list from SNR and linear weighted average methods were input into GeneGo pathway analysis program (www.genego.com). The Gene ontology process/pathway list was generated using a cut-off of $p < 0.05$ for each analysis. From those lists, the common pathway list was generated.

[b]Pathways that are part of—or overlap with—the testosterone synthesis pathways are indicated by bold italics. These pathways were identified by performing a PubMed literature search (http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed) for "testosterone" and the name of each pathway (listed in the table).

[c]Entrez Gene indicates that *Insl3* is the ligand for the LGR8 receptor, but the *Insl3* pathway is not fully defined (http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=114215&ordinalpos=3&itool=EntrezSystem2.PEntrez.Gene.Gene_ResultsPanel.Gene_RVDocSum). Functions that have been shown to be related to the *Insl3* pathway are G-protein-coupled receptor binding and hormone activity. Processes identified are G-protein signaling, adenylate cyclase inhibiting pathway, gonad development, in utero embryonic development, male gonad development, negative regulation of apoptosis, negative regulation of cell proliferation, oocyte maturation, positive regulation of cAMP biosynthetic process, and positive regulation of cell proliferation. While a number of G-protein pathways were identified in this analysis, none are considered exclusive to *Insl3* and are, therefore, not listed in bold italics.

Table 6-4. Genes involved in cholesterol biosynthesis/metabolism as identified by the two analyses (i.e., linear weighted normalization and signal to noise ratio) of Liu et al. (2005)

Linear weighted normalization (GeneGo)	SNR (GeneGo)	SNR (KEGG)
	Acat1	*Acat1*
Cyp27a1		
Cyp51a1	*Cyp51a1*	
Cyp7b1		
Dhcr7	*Dhcr7*	*Dhcr7*
	Dhcr24	
	Ebp	*Ebp*
	Fdft1	*Fdft1*
	Fdps	*Fdps*
Hmgcr	*Hmgcr*	*Hmgcr*
Hmgcs1	*Hmgcs1*	*Hmgcs1*
Hsd11b1		
Hsd3b1		
Idi1	*Idi1*	*Idi1*
	Mvd	*Mvd*
	Nsdhl	
Sqle	*Sqle*	*Sqle*
Sc4mol	*Sc4mol*	
Soat1		
	Tm7sf2	

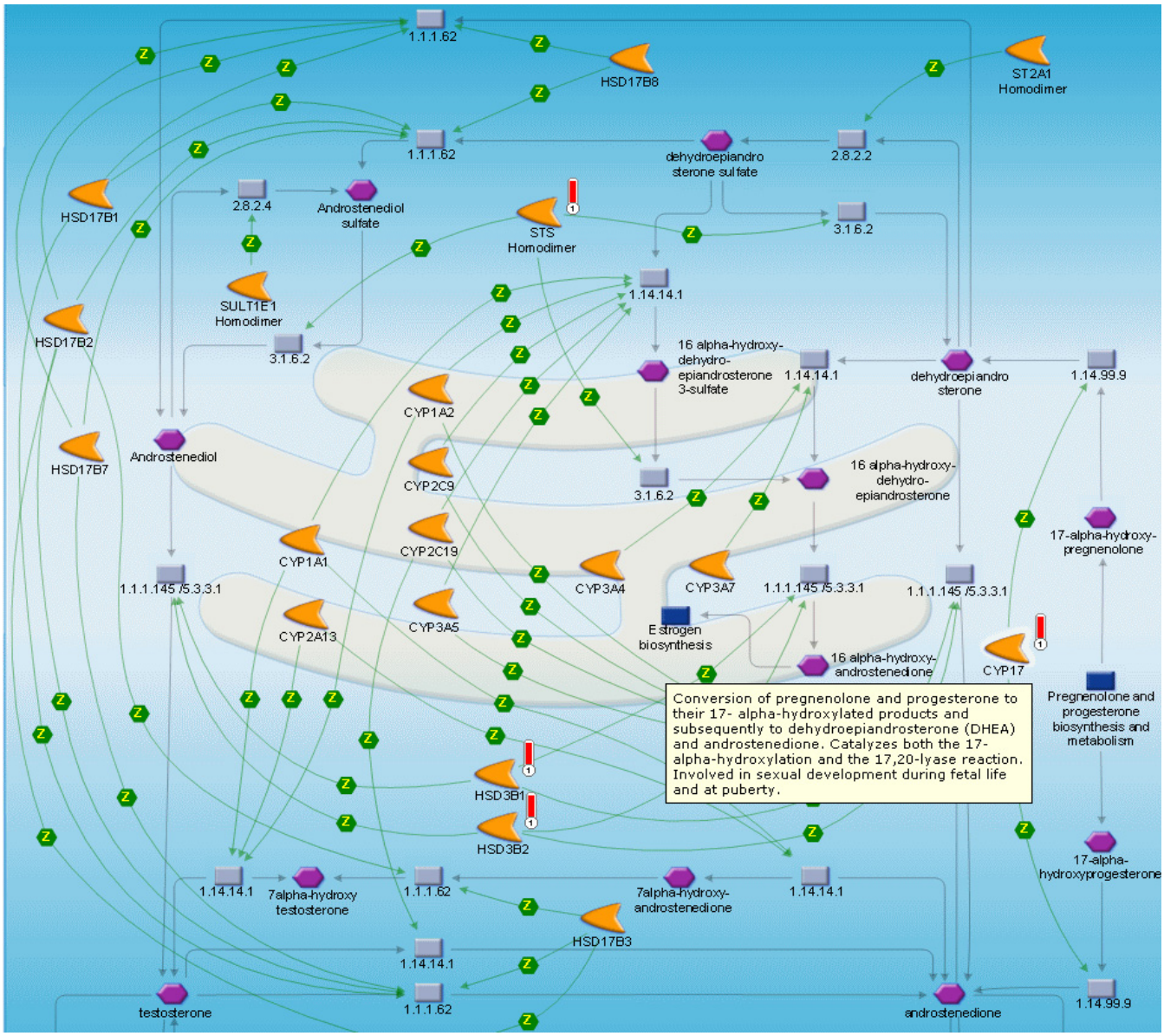

Figure 6-5. Mapping the Liu et al. (2005) data set onto the canonical androstenedione and testosterone (T) biosynthesis and metabolism pathway in MetaCore™ (GeneGo). Key enzymes activated by DBP are identified by red thermometers.

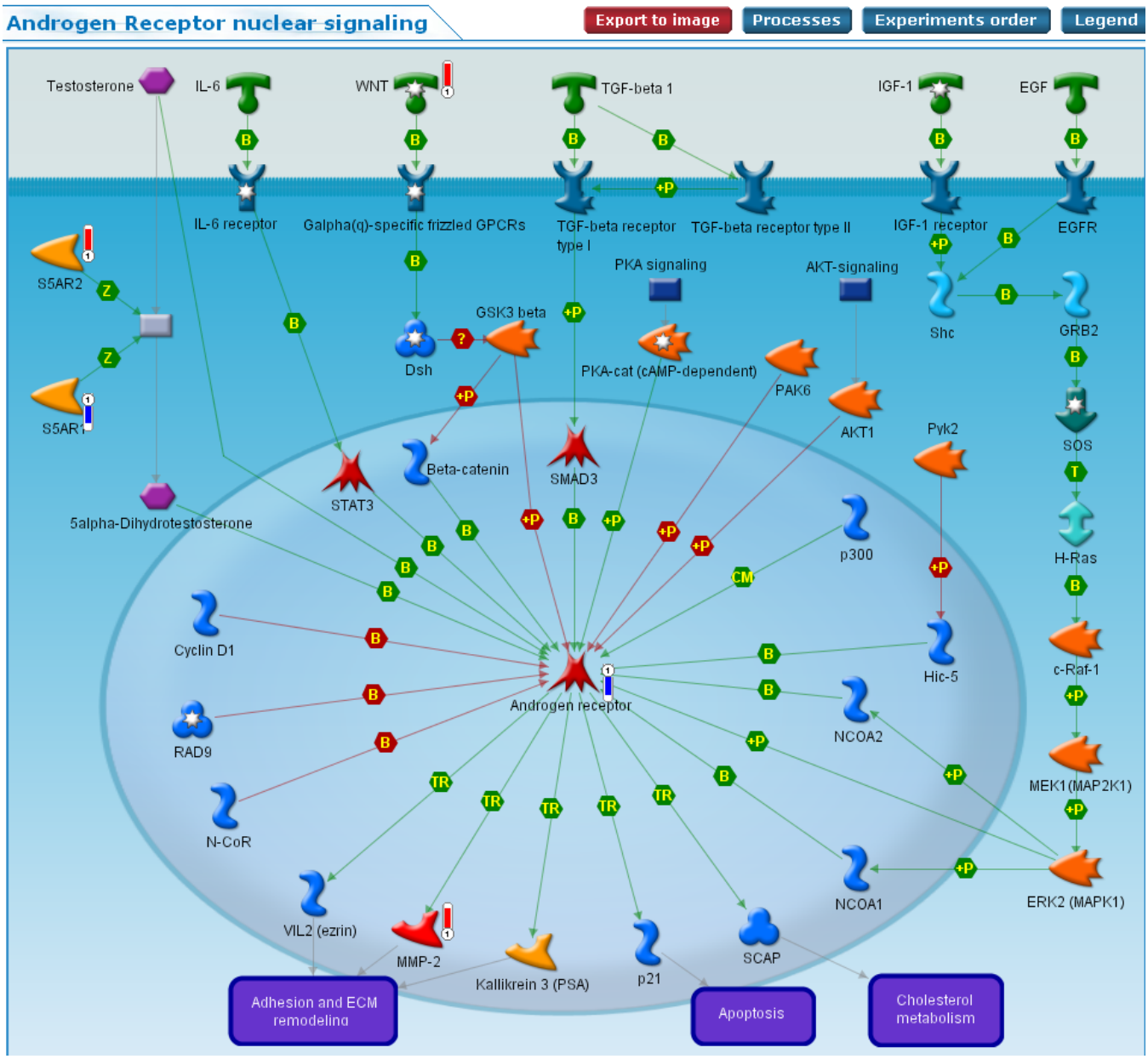

Figure 6-6. Mapping the Liu et al. (2005) data set onto the canonical Androgen receptor (AR) nuclear signaling pathway in MetaCore™ (GeneGo). The thermometers denote input intensities of genes from our statistical list mapped to this GeneGo pathway. Blue thermometers represent downregulated genes present in the data and red thermometer represents upregulated genes present in the data set that map to this pathway.

The GeneGo network connections reveal that CYP17 and AR are involved in the androgen biosynthetic process. Based on the transcriptional profiling data, the AR is down

regulated by DBP in the fetal testes. This was a novel finding from this analysis and needs further corroboration.

It has been reported in the literature (MAQCI, see Chapter 2) that the results of a microarray experiment are heavily dependent on the data analysis protocol and the biological pathway analysis tools available to interpret the list of statistically significant genes. Dissimilar sets of gene expression signatures with distinct biological contexts can be generated from the same raw data by different data analysis protocols. Distinct biological contexts can also be generated from the same gene expression signatures by different biological pathway protocols. Therefore, it becomes important to determine and understand the relationship between the gene expression and pathway changes and a biological outcome of interest.

To do a thorough investigation it is necessary to use many sources of gene and pathway annotation. The intent of using multiple sources is to gain an enriched analysis. In practice, analysis is carried out with the suite of tools available to the analyst. In this case, the Star Center primarily used KEGG (a resource rich in enzymatic and metabolic reactions but weak in signaling pathways); whereas the U.S. EPA used Rosetta Resolver, GeneGo, and Ingenuity Pathway Analysis, resources that are populated with signaling as well as metabolic pathways.

This exercise demonstrates that multiple approaches to microarray data analysis can yield similar biologically relevant outcomes. The differences observed in the results could be due to a number of factors including (1) the different data normalization procedures used in the two separate analyses; (2) different data interpretation tools such as the software for pathway analyses, for examples. However, it cannot be ruled out that the differences may reflect differences in biological significance (i.e., one approach is better than the other).

6.2.3. Transcription Factor (TF) Analysis

Inspection of the regulatory elements of the informative genes would reveal important information about DBP exposure on gene expression. All the informative genes demonstrated a down regulation in expression, and their co-regulated genes are likely to have a similar response (Turner et al., 2007). EXPANDER is used for TF enrichment analysis (Shamir, 2005). TF enrichment analysis revealed six transcription factors in informative genes with a statistical significance level of 0.05 (see Table 6-5). Liu et al. (2005) study states that the regulatory regions of several steroidogenic genes contain Globin transcription factor 1 binding protein

(GATA) elements, and propose that GATA factors, particularly GATA-4 and GATA-6, might represent novel downstream effectors of hormonal signaling in steroidogenic tissues. Interestingly, GATA-4 appears in the DEG list as up regulated, and GATA-1 is one of the enriched transcription factors. Another study claims that estrogen receptor (ER) α-deficient mice (ERα-/-) display higher levels of testicular T secretion than wild-type mice from fetal day 13.5 onwards (Delbes et al., 2005) and that ER is expressed in the rat testis (van Pelt et al., 1999). Sex determining region Y (SRY) is one of the enriched transcription factors. Although SRY is known to be the major determinant for testis formation, a recent study showed that SRY is expressed also in rat testis tissues (Turner et al., 2007). Nuclear factor Y (NF-Y) is another putative transcription factor, and it is known as taking action in sterol regulation (Shea-Eaton et al., 2001; Xiong et al., 2000).

Table 6-5. Enriched transcription factors (TFs) from Liu et al. (2005) data set

Transcription factor[a]	*p*-Value[b]
ER	0.00297
GATA-1	0.00966
AREB6	0.0197
SRY	0.0385
NF-Y	0.0407
Nrf2	0.0462

[a]PRIMA (Promoter Integration in Microarray Analysis) is used to identify transcription factors whose binding sites are enriched in a given set of genes promoter regions.
[b]The enrichment score of the transcription factors: $p < 0.05$ cutoff.

6.3. DEVELOPMENT OF A NEW METHOD FOR PATHWAY ANALYSIS AND GENE INTERACTIONS: PATHWAY ACTIVITY LEVEL (PAL) APPROACH

An alternative approach to infer important biological pathways is based on the use of the available knowledge of functional annotations prior to statistical analysis. Based on the assumption that the expression levels of sets of genes that are functionally related follow similar trajectories, due to activation or deactivation of a pathway under different environmental

conditions or at different time points, average correlation between genes in a given pathway leads to significant findings (Kurhekar et al., 2002; Pavlidis et al., 2002; Zien et al., 2000). It is not required that all the genes follow the same pattern. In these pathway scoring methods, a pre-defined cut-off value is applied to determine the number of genes to be included. However, focusing only on genes with a pre-determined significance analysis in gene expression may result in a loss of information.

An alternative method to define the systematic behavior of the pathway is to evaluate the pathway activity level (PAL), the method suggested by (Tomfohr et al., 2005). The strength of the PAL method over other pathway analyses is that the expressions of all genes within a pathway are considered.

The procedure begins with mapping genes to the KEGG pathway database. The entire gene set represented by the Liu et al. (2005) data set (i.e., using the Affymetrix® RAE230 A and B chips) maps to 168 pathways in the KEGG database with 2,483 associated genes. Gene expressions are z-scored before the analysis. Using Equation 6-2, let $\Xi_{p(k,t)}$ be the gene expression matrix of a given pathway p of size k genes and t arrays (i.e., t-different time points). Tabulate the normalized (i.e., to zero mean and a unity standard deviation) gene expression data. Each element of $\Xi_{p(k,t)}$ is the relative expression level of the k^{th} gene in the t^{th} time point. The vector in the k^{th} row of the matrix $\Xi_{p(k,t)}$ lists the relative expression of the k^{th} gene across the different time points.

$$\Xi_{p(k,t)} = U_p(k,k) \times S_p(k,t) \times V_p(t,t)^I \qquad (6\text{-}2)$$

Equation 6-2 states that the matrix $\Xi_{p(k,t)}$ can be decomposed to a rotation matrix, $U_p(k,k)$, a stretch matrix, $S_p(k,t)$, and a second rotation matrix, $V_p(t,t)$. $U_p(k,k)$ is an orthonormal basis that spans the gene expression space of Ξ_p, whereas $V_p(t,t)$ is an orthonormal basis spanning the sample (array) space of $\Xi_{p(k,t)}$, that forms a set of new basis vectors for the columns of $\Xi_{p(k,t)}$. $S_p(k,t)$ is a diagonal matrix (i.e., eigenvalue matrix), whose elements are sorted from highest to the lowest based on the magnitude of the singular values. In Equation 6-3, the PAL of a given pathway is defined as the projection onto the first eigenvector

that spans the sample (array) space $\Xi_{p(k,t)}$. Thus, gene expression levels are reduced to pathway activity levels.

$$PAL_p(n) = V_p(n,1)^I \tag{6-3}$$

$PAL_p(n)$ is a *1xn* vector, and each entry represents the pathway activity level of corresponding sample. If n_1 samples are denoted as control experiments and n_2 have undergone some type of treatment then the activity levels are given in Equations 6-4 and 6-5.

$$PAL_1(p) = V_p(n_1,1)^I \tag{6-4}$$

$$PAL_2(p) = V_p(n_2,1)^I \tag{6-5}$$

Activity levels represent the cumulative effect of gene expressions in a given pathway and therefore the relative activity. The next step is to quantify the differentiation between pathway activities of the treatment groups, control and treated. Overall pathway activity (OPA) denotes the change of pathway activity levels between different groups (Equation 6-6). For a given pathway p:

$$OPA_p = \frac{\left|\mu(PAL_1) - \mu(PAL_2)\right|}{\sigma(PAL_1) + \sigma(PAL_2)} \tag{6-6}$$

μ and σ denote the mean and standard deviation of the activity levels, as evaluated using Equation 6-6 for pathway *p*. A higher OPA indicates a better discrimination between pathway activity levels of vehicle and treated samples. To compute the statistical significance of

the OPA of a given pathway, we randomly permute the gene expression data on the chip for each pathway and calculate the pathway activity levels and OPA 1000 times via Equations 6-4, 6-5, and 6-6. If the fraction of the artificial OPA that are higher than the actual OPA exceeds 0.05, or any other appropriately defined statistical significance level, the actual OPA is attributed to random variations. The pathways that exhibit statistically significant high OPA are defined as "active pathways." Appendix B shows the algorithm for selecting statistically significant pathways (see Figure B-2). This calculation allows us to rank, and compare, active pathways based on their OPA. The term "active pathway" does not indicate any up-regulation or down-regulation, but rather indicates an overall change of the pathway compared to control samples. Thus, an "active" pathway can still be one that is reduced and nonfunctional following chemical treatment, and an "inactive pathway" can still be functional but not exhibit significant difference from the control. Of the 168 KEGG pathways that mapped to the Liu et al. (2005) data set, only 32 were found to be active pathways with an OPA level of less than $p = 0.05$ (see Table 6-6).

This analysis identified valine, leucine, isoleucine (VLI) degradation, sterol biosynthesis, citrate cycle, and fatty acid metabolism as the most active pathways due to DBP exposure. Figure 6-7 depicts the active pathways and their connections via metabolites, from the most active pathways towards the least active pathways based on OPA. The connections of the active pathways are retrieved from KEGG. The statistical outcome of the pathway activity analysis and the relationship between active pathways are integrated. The active pathways have connections to non-active pathways; but only active pathways are included in the metabolic network. It is shown that the active pathways identified in this study are linked together at the metabolite level indicating biological significance.

Table 6-6. Statistically significant pathways as derived by signal-to-noise ratio analysis[a]

Pathway[b]	Activity	*p*-value[c]
Inositol metabolism	1.6338	0.0328
Reductive carboxylate cycle CO_2 fixation	1.65	0.0444
Galactose metabolism	1.7422	0.0475
Pentose phosphate pathway	1.8216	0.0467
Pyruvate metabolism	1.8747	0.0435
Glycolysis/gluconeogenesis	2.0128	0.0456
Fructose and mannose metabolism	2.1187	0.0405
Pentose and glucuronate interconversions	2.1545	0.0315
Carbon fixation	2.2202	0.0337
Synthesis and degradation of ketone bodies	2.2333	0.0224
Alanine and aspartate metabolism	2.4667	0.0235
Phenylalanine metabolism	2.4877	0.0212
Propanoate metabolism	2.5783	0.0224
Citrate cycle TCA cycle	2.6658	0.0218
Benzoate degradation via CoA ligation	2.6678	0.0145
C21-Steroid hormone metabolism	2.911	0.0136
Metabolism of xenobiotics by cytochrome P450	3.0373	0.0245
Tryptophan metabolism	3.0424	0.0205
Ascorbate and aldarate metabolism	3.1052	0.0095
Glutathione metabolism	3.1356	0.0182
Terpenoid biosynthesis	3.3621	0.0044
Lysine degradation	3.4557	0.0121
Fatty acid metabolism	3.4732	0.0154
Limonene and pinene degradation	3.4945	0.0072
Arginine and proline metabolism	3.7056	0.011
Histidine metabolism	3.71	0.0084
Glycine, serine and threonine metabolism	3.9578	0.0092
beta-alanine metabolism	4.1212	0.0063
Butanoate metabolism	5.1243	0.0023
Biosynthesis of steroids	5.3459	0.0011
Valine, leucine and isoleucine degradation	5.6232	0.003
Alkaloid biosynthesis	5.6922	0.001

[a]Pathways: that are defined in KEGG.
[b]Activity: quantifies the difference between different experimental conditions (i.e., corn oil control and DBP-treated samples).
[c]Significance analysis of activities: $p < 0.05$ cutoff for significant pathways perturbed by DBP exposure.

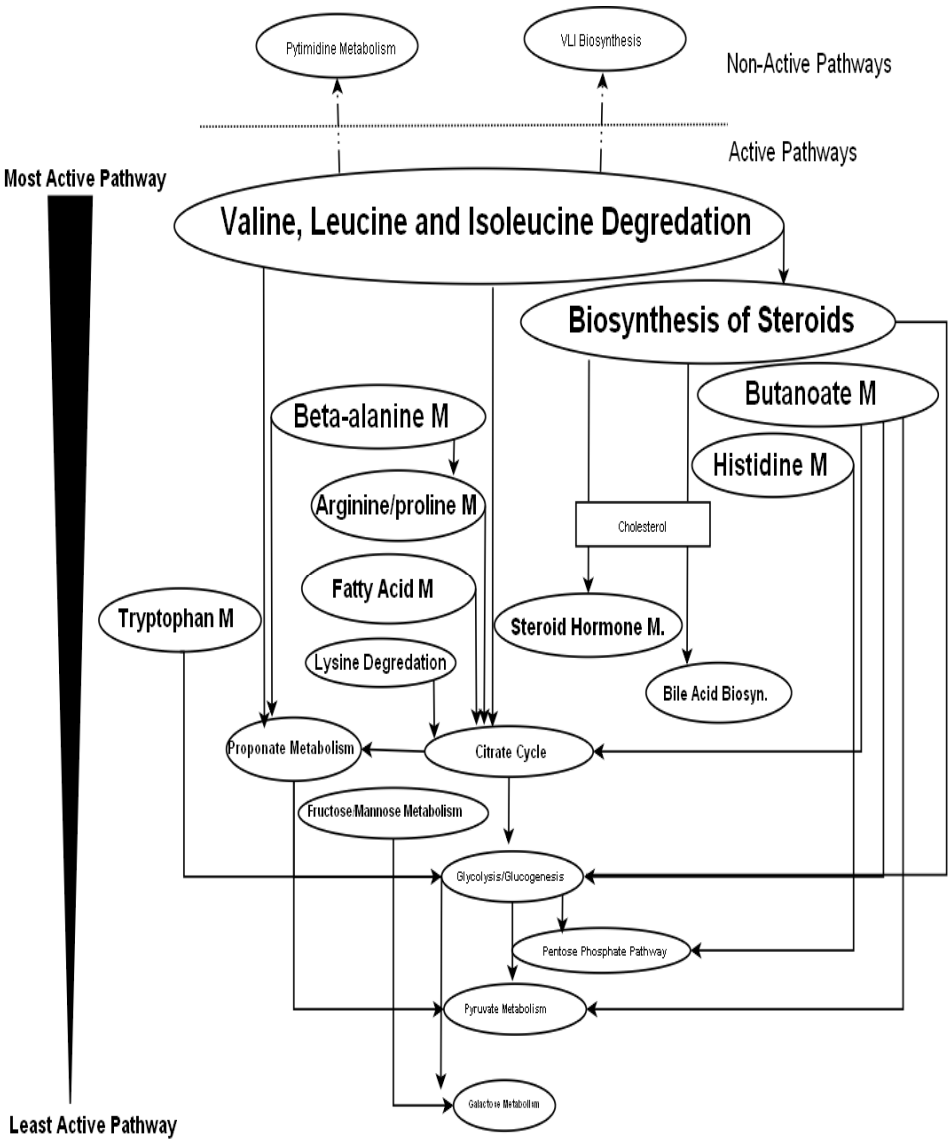

Figure 6-7. Statistically significant pathway interactions generated using the KEGG database following overall pathway activity (OPA) analysis. The Liu et al. (2005) data set used for analysis. ⬭ = pathway, ▭ = metabolite. Larger oval sizes indicate relative impact on a pathway, where the larger ovals indicate a greater effect on a pathway after DBP exposure.

The value of this approach depends on the content of the employed pathway database. For example, some of the pathways may not be present in testes tissue. For example, even though bile acid biosynthesis does not occur in the testis, the collection of genes related to bile acid biosynthesis showed statistically significant change.

OPA is a linear projection of gene expressions that constitute a given pathway. The singular value decomposition analysis quantifies the variance between two experimental groups in the context of the pathway. To examine the effect of statistical significance of gene expression on a given pathway, the OPA of a pathway is calculated by adding genes one at a time starting with the gene with the highest SNR. Subsequently, the next gene with the second highest SNR in this pathway is identified and added, etc., until all genes in the pathway have been used to determine the OPA**.** Then, the Equations 6-4, 6-5, and 6-6 (section 6.3.) are reevaluated with two genes and so forth until all of the genes in the given pathway have been included. Figure 6-8 illustrates an example of this process for determining active and inactive pathways, evaluating the Liu et al. (2005) DBP data. The inactive mTOR pathway has only a single gene with a high SNR. As additional genes within the pathway with much lower SNR are considered, the OPA is reduced. In contrast, the active VLI degradation pathway has numerous genes with high SNR, and as all genes within the pathway are considered, the OPA remains high. From this analysis, we determined that there is a subset of genes with high SNR that maintain the OPA score for active pathways. We define DEGs that are in active pathways as informative genes (see Table B-1). The interactions between informative genes were retrieved via IPA® and the resulting preliminary gene network is shown in Figure 6-9.

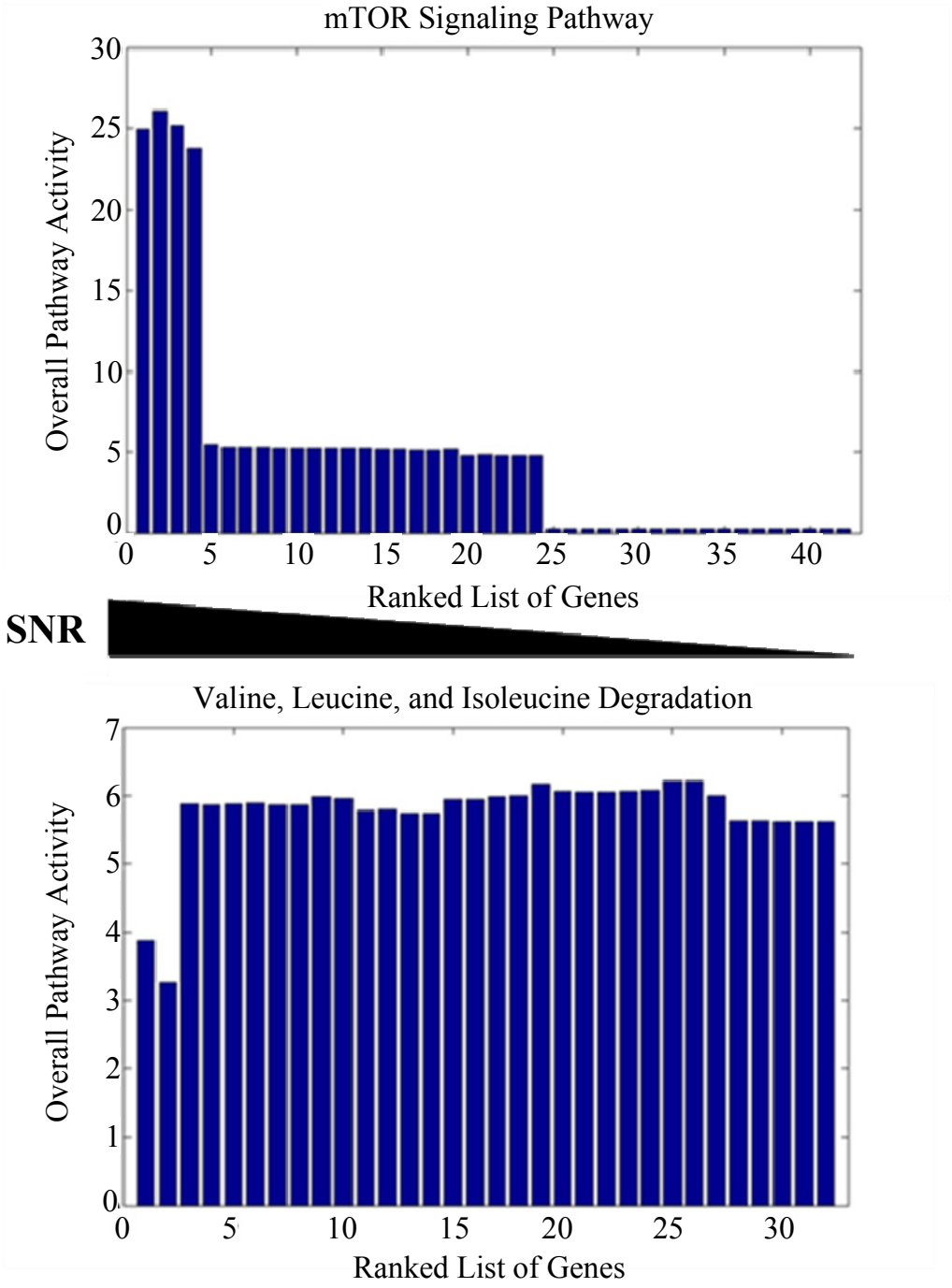

Figure 6-8. Overall pathway activity (OPA) of the affected pathways calculated by adding genes according to the decreasing signal-to-noise ratio (SNR). The Liu et al. (2005) DBP only data were evaluated using the OPA method.

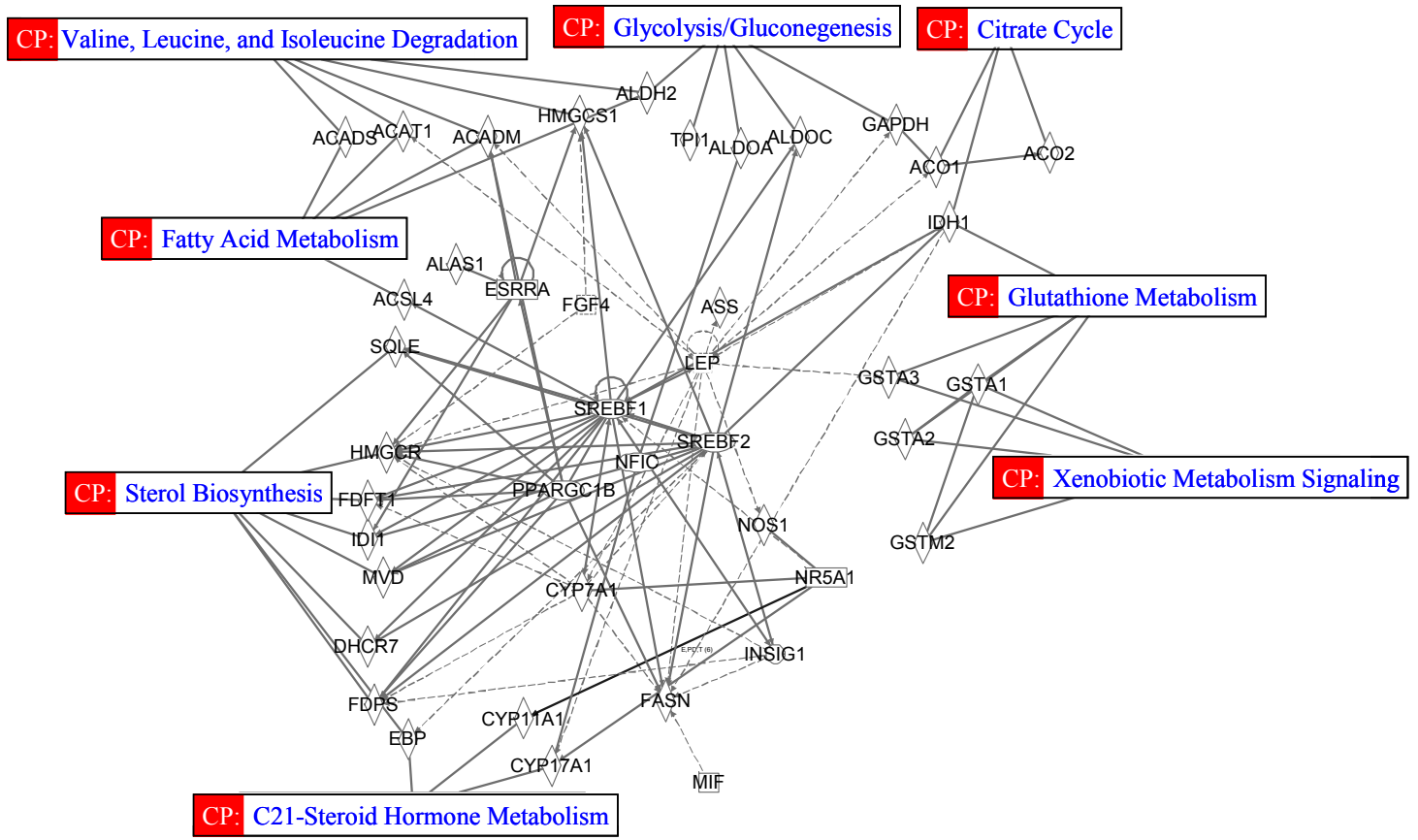

Figure 6-9. Gene network after DBP exposure created by Ingenuity® Pathway Analysis (IPA) from the informative gene list. This model is based on data from Liu et al. (2005). This model illustrates the interactions among genes after DBP in utero exposure in the rat testis. Genes (noted in Table B-3) are added in from the Ingenuity® knowledgebase. . Active pathways, which do not share any common metabolites with other active pathways, may interact via added nodes and informative genes. Genes or gene products are represented as nodes. Diamonds, enzymes; Horizontal ovals, transcription regulators; Squares, cytokines; Rectangles, nuclear receptors; Solid lines, direct relationship between edges (i.e., 2 nodes; 2 molecules that make physical contact with each other such as binding or phosphorylation); Dashed lines, indirect interactions (i.e., do not require physical contact between the two molecules, such as signaling events) between edges.

6.4. EXPLORING GENETIC REGULATORY NETWORK MODELING: METHODS AND THE DBP CASE STUDY

The goal was to utilize existing DBP genomic data to develop a regulatory network model useful to risk assessment. Genetic regulatory network models illustrate interactions between genes and their products (e.g., mRNA, proteins). Network models encompass identified pathways from input data and in addition incorporate gene elements that are inferred from the input data. The availability of one time course study data enabled us to model the series of events that occurred between exposure to DBP and the onset of adverse reproductive outcomes by the generation of a regulatory network model. We used Ingenuity® Pathway Analysis (IPA) software to identify the relationships among the informative genes. IPA adds nodes (i.e., genes) to the input gene list (i.e., informative genes) and then, builds edges (i.e., relations) based on the literature to develop a regulatory network (Sladek et al., 1997).

Time-course studies are ideal for developing regulatory network models of biological processes to model the dynamic networks for formulating mechanistic explanations of dynamic developmental mechanisms. The Thompson et al. (2005) study was selected because it was the only study that had time-course data. Additionally the study had the advantage of using the Affymetrix® chip, which has ~30,000 rat genes represented, and the data were provided by Dr. Kevin Gaido, one of our collaborators. Thompson et al. (2005) conducted a study where animals were exposed to DBP for 30 minutes and 1, 2, 3, 6, 12, 18, and 24 hours on GD 18 and 19. The limitations of the Thompson et al. (2005) study include: 1) the dosing was initiated on GD 18, quite late in the critical window, and 2) the shortest duration exposure began at the latest developmental time (i.e., duration and developmental stage do not coincide; see Chapter 5). Given this caveat, the data were utilized because it was the only study available to test algorithms to build a prototype of a regulatory network model.

We used the PAL method, described earlier, to identify biologically active pathways at each time point. We evaluated the informative genes at each time point and the resulting preliminary gene network, based on the Thompson et al. (2005) data, is shown in Figure B-3. The analysis showed a preponderance of signaling pathways such as JAK/STAT, PPAR, and MAPK perturbed at the earlier exposure durations with the metabolic pathways being affected following longest exposures to DBP (18 hours). The majority of the active pathways at this

dose-exposure time (18 Hour) are metabolic pathways such as amino acid metabolism, lipid metabolism, and carbohydrate metabolism. Thompson et al. (2005) hypothesized that the decrease in T level after a short duration of DBP exposure might be because of the cholesterol unavailability. Their study findings support this hypothesis. To have a complete understanding of the temporality of the DBP effect, data from an exposure-duration series across the entire critical window of exposure are needed.

6.5. EXPLORING METHODS TO MEASURE INTERSPECIES (RAT TO HUMAN) DIFFERENCES IN MOA

The goal was to address Case Study Question 2, whether genomic and mechanistic data could inform the interspecies (rat to human) differences in MOA, was explored. Although progress has been made over the past four decades in understanding the MOA of chemical toxicants, it is increasingly important to determine mechanistically the relevance of these MOAs in humans. With the sequencing of the human, mouse, and rat genomes and knowledge of cross species gene and protein homologies, the studies of differential gene expression in animal models have the potential to greatly enhance our understanding of human disease. Genes co-expressed across multiple species are most likely to have conserved function. The rat genome project reported that almost all human genes known to be associated with disease have orthologous genes in the rat genome, and that the human, mouse, and rat genomes are approximately 90% homologous (Gibbs et al., 2004). Because the function of a specific gene and its involvement in disease might not be conserved across species, along with structural and functional homology, the conservation of function of blocks of genes—i.e., pathways—are likely to be more important in cross species comparison (Fang et al., 2005).

In the absence of DBP genomic data in human cell lines, we considered genetic sequence data as a source of genomic data for making species comparisons. Even if such data were available, in vivo (rat genomic data) to in vitro (human genomic data) extrapolations may confound the ability to generate an accurate interspecies comparison. Use of bioinformatic approaches to examine microarray expression profiles from exposure to a chemical in an animal model to elucidating genes and pathways that might be associated with exposure in humans holds great promise. Similarity analysis between single gene and protein sequence analysis cannot represent the complex relationships species therefore species comparison studies emerged

to compare pathways to analyze a higher level organization. Attempts include reaction content (Hong et al., 2004), enzyme presence (Heymans, 2003), and enzyme sequence information of the enzymes in a given pathway (Forst et al., 1999, 2001). The pathways for the biosynthesis of steroids have a lot of similarity between humans and rats. Protein sequence similarity, cross-species pathway network similarities, and promoter region conservation cross-species comparisons to evaluate cross-species similarity metrics were performed. The results from comparing the predicted amino acid sequence similarities between rat and human for the steroidogenesis pathway proteins are shown in Table 6-7.

Table 6-7. The enzyme sequence similarity of the enzymes of steroidogenesis pathway between rat and human

Gene symbol	Entrez gene ID	mRNA and protein IDs	Human homolog IDs	Identities[a]	Positives[b]	Gaps[c]
Dhcr7	64191	NM_022389.2→NP_071784.1	Q9UBM7	412/475 (86%)	443/475 (93%)	4/475 (0%)
Idi1	89784	NM_053539.1→NP_445991.1	AF003835	196/227 (86%)	215/227 (94%)	0/227 (0%)
Fdps	83791	NM_031840.1→NP_114028.1	M34477	301/353 (85%)	326/353 (92%)	0/353 (0%)
Fdft1	29580	NM_019238.2→NP_062111.1	AAP36671	356/413 (86%)	393/413 (95%)	0/413 (0%)
Hmgcr	25675	NM_013134.2→NP_037266.2	AAH33692	738/890 (82%)	768/890 (86%)	58/890 (6%)
Mvd	81726	NM_031062.1→NP_112324.1	AAP36301	338/398 (84%)	357/398 (89%)	1/398 (0%)
Sqle	29230	NM_017136.1→NP_058832.1	NP_003120	481/574 (83%)	528/574 (91%)	1/574 (0%)
Ebp	117278	NM_057137.1→NP_476478.1	NP_002331	618/732 (84%)	673/732 (91%)	1/732 (0%)
Lss	81681	NM_031049.1→NP_112311.1	NP_002331	618/732 (84%)	673/732 (91%)	1/732 (0%)
Sc5d	114100	NM_053642.2→NP_446094.1	NP_008849	246/299 (82%)	275/299 (91%)	0/299 (0%)
Mvk	81727	NM_031063.1→NP_112325.1	BAD92959	323/393 (82%)	355/393 (90%)	0/393 (0%)
Cyp27b1	114700	NM_053763.1→NP_446215.1	NP_000776	413/508 (81%)	453/508 (89%)	7/508 (1%)
Nqo1	24314	NM_017000.2→NP_058696.2	NP_000894	234/274 (85%)	250/274 (91%)	0/274 (0%)
Vkorc1	309004	NM_203335.2→NP_976080.1	AAQ13668	83/94 (88%)	88/94 (93%)	0/94 (0%)
Average similarity scores				84%	94.14%	

Table 6-7. (continued)

[a]Identities: The number and fraction of total residues in the HSP which are identical.
[b]Positive: The number and fraction of residues for which the alignment scores have positive values.
[c]Gap: a space introduced into an alignment to compensate for insertions and deletions in one sequence relative to another. To prevent the accumulation of too many gaps in an alignment, introduction of a gap causes the deduction of a fixed amount (the gap score) from the alignment score. Extension of the gap to encompass additional nucleotides or amino acid is also penalized in the scoring of an alignment.

The HSP (high-scoring segment pair) is the fundamental unit of BLAST algorithm output. Alignment: The process of lining up two or more sequences to achieve maximal levels of identity (and conservation, in the case of amino acid sequences) for the purpose of assessing the degree of similarity and the possibility of homology.

Source: http://searchlauncher.bcm.tmc.edu/help/BLASToutput.html#anchor14684156.

Our analysis suggests that the biosynthesis of steroids is highly conserved across humans and rats, with the average sequence similarity of enzymes between human and rat being ~87% and the average promoter region conservation of genes at 52% (see Table 6-7). However, it is difficult to unequivocally determine a "high" versus "low" degree of conservation for the genes in this pathway—especially in light of the fact that the more important gene products (such as a rate-limiting step) have not been identified for DBP on steroidogenesis. Additionally, there are likely differences between a statistically meaningful "high" degree of conservation vs. an understanding of whether the biologically meaningful regions of the predicted protein sequence are conserved.

Cross-species pathway network comparison is a creative approach using network data from publicly available databases to assess species similarities. However, uncertainties and gaps in the database information at this time make conclusions difficult. Therefore, these data are not described herein.

Development of new bioinformatic and statistical resources using data generated in human cell lines, together with the information obtained from rat in vivo studies may provide new, useful data to further investigate interspecies differences in response to a chemical agent. To determine the viability of using such metrics to inform the interspecies concordance of mechanism issue in risk assessment, homology-based analysis of genes and proteins need to be conducted in systems where the concordance in mechanism across species is well established by prior studies to serve as a base line for "high homology."

6.6. CONCLUSIONS

The projects to address the four objectives presented in this chapter serve as a broad range of examples of genomic data analyses available to the risk assessor with expertise (or collaborators with expertise) in bioinformatics, and in some cases, represent exploratory efforts to develop methods for analyzing genomic data for use in risk assessment. These methods include DEG identification, pathway level analysis (including the newly described OPA method), regulatory network analysis, and tools to assess cross-species similarities in pathways. A summary for a less technical audience than the remainder of this chapter is presented next, grouped by the four subobjectives for the work.

- Reanalyze DBP microarray data to address the Case Study Question 1: *Do the genomic data inform DBP additional MOAs and the mechanism of action for the male reproductive developmental effects?*

  We performed a number of reanalyses of the Liu et al. (2005) data because the pathway analysis presented in the article was not performed for risk assessment purposes. While the authors of this and other microarray studies support two MOAs for DBP, a reduction of fetal testicular testosterone via affects on steroidogenesis and cholesterol transport genes, not all pathways associated with the differentially expressed genes were discussed in detail.

  Two different bioinformatics tools to analyze the same data were compared. Each analysis used multiple statistical filters to parse the noise from the signal in the microarray data set and to assess the quality of the data set. Ideally, for a high quality study data set, there would be a minimum of variance between similarly treated samples and the variance would lie between the control and treated sample data. PCA shows the quality of the Liu et al. (2005) data set to be of moderate quality based on the observed variance among similarly treated data sets (control and treated groups). One analysis utilized multiple proprietary software packages (GeneGo, Rosetta Resolver). The rationale for looking at the effect of DBP on the pathway level as opposed to a cluster of genes is that DBP is most likely affecting multiple pathways within a cellular environment. The methods comparison exercise allowed us to generate a list of affected pathways in common between the two methods, and in this way, provided more confidence focusing on these pathways.

  The results of the new pathway analyses both corroborate the previously identified two MOAs for DBP male reproductive development toxicity, and provide putative novel pathways affected by in utero DBP exposure that may play a role in DBP-mediated toxicity. The results of the new pathway analyses provide hypotheses for MOA that could be tested in new experimental studies. Future research could investigate the role of these pathways in DBP-induced toxicity. In addition, a gene network was developed for DBP based on the Liu et al. (2005) data. The GeneGo analysis and the validating the role of the steroidogenesis pathway also revealed the modulation in CYP17 and AR that are involved in the androgen biosynthetic process. This is a new hypothesis that requires followup with new studies to confirm this observation. Performing new analyses was useful for the purposes to further our understanding of the DBP mechanism of action.

- *Explore the development of new methods for pathway analysis of microarray data for application to risk assessment.*

  Quality control requirements for microarray study analysis for use in risk assessment are distinct from their use in basic research. In traditional pathway level analysis, significant genes are mapped to their respective pathways. Depending on whether the number of

genes that map to any given pathway, the role of the pathway can be over of underestimated. To overcome this problem, we developed the overall pathway activity (OPA) method that employs one as opposed to two steps (i.e., first, identifying DEGs and second, identifying significantly affected pathways by grouping the DEGs using pathway analysis programs). This method scores pathways based on the expression level of all genes in a given pathway.

The OPA analysis identified valine, leucine, isoleucine (VL1) degradation, sterol biosynthesis, citrate cycle, and fatty acid metabolism as the most active pathways following DBP exposure. These findings support the hypothesis of Thompson et al. (2005), that an early decrease in testosterone levels may be a result of cholesterol unavailability. However, for this approach to be useful, knowledge of tissue-specific pathways is required. For example, even though bile acid biosynthesis does not take place in the testis, a pathway related to bile acid biosynthesis was identified as statistically significant in this analysis. Further developed on the OPA method needs to incorporate tissue-specific relevant. This method shows promise for use in risk assessment.

- *Utilize existing DBP genomic data to develop a genetic regulatory network model, and methods for modeling, for use in risk assessment.*

Genetic regulatory network models can be very useful for understanding the temporal sequence of critical biological events perturbed after chemical exposure, and thus, useful to a risk assessment. We developed a method for developing a genetic regulatory network model for DBP based on the available data. The availability of a time-course data (Thompson et al. [2005]) enabled our group to model the series of events that occurred between exposure to DBP and the onset of toxic reproductive outcomes by the generation of a regulatory network model. However, given the limitations of the Thompson et al. (2005) study design, we did not draw conclusions about affected genes and pathways over time for DBP from this study. Instead, the Thompson et al. (2005) data was used to build a prototype of a regulatory network model and thus, the exercise allowed us to develop methods for analyzing time course data for use in building a regulatory network model.

- *Utilize genomic and other molecular data to address the Case Study Question 2: Do the genomic and other molecular data inform interspecies differences in MOA?*

Extrapolation from animal to human data is critical for establishing human relevance of an MOA in risk assessment. Genes co-expressed across multiple species could have a conserved function. The human, mouse, and rat genomes have been reported to be 90% homologous (Gibbs et a., 2004). However, because it is not certain whether the function of a specific gene is conserved across species, conservation of pathways across species can be one important factor in establishing cross species concordance of MOA. In

addition, a common critical role of androgens in both rodent and human male development of reproductive organs has been well established.

Using the available DNA, sequence, and protein similarity data for the steroidogenesis pathway, we used three different methods to assess rat-to-human conservation as metrics that may inform the interspecies differences for one MOA, the reduced fetal testicular T. The pathways for the biosynthesis of steroids have similarity between humans and rats. Comparing the predicted amino acid sequences for the steroidogenesis pathway genes, we found that the average sequence similarity between rat and human is ~87% and the average promoter region similarity of genes is 52%. Some of the challenges in using similarity scores to estimate the cross species relevance of a MOA are described (section 6.5.).

In summary, the preliminary analytical efforts described in this chapter address and raise a number of issues about the analysis of microarray data for risk assessment purposes. First, analyzing any given data set multiple ways and arriving at the same conclusion provides confidence in the analytical approach—however, there is no "gold standard" analytical method. Second, applying stringent statistical filters in pathway analysis (e.g., $p < 0.05$, Benjamini Hochberg multiple testing correction) can limit the number of genes that are identified. Interpretation of the biology of the system using only a limited gene set is restrictive. It is important to remember that the genes that do not pass the statistical stringency cut-off may be crucial for understanding the biology of the system, as statistical significance and biological significance are not necessarily the same. Therefore, it becomes incumbent upon the researcher, to analyze the data in multiple ways in order to maximize the benefits of this technology. Third, a pathway level analysis restricts the incorporation of all genes for determining relevant pathways that are affected by DBP. There is a substantial amount of background noise generated in a typical microarray experiment (i.e., gene expression variability even among the controls; see Smith, 2000). For use in risk assessment, it is important to be able to identify and separate the signal from the noise. Innovative approaches such as the OPA method described in this chapter may provide more confidence when evaluating microarray data for use in risk assessment. These efforts reveal some of the promises and challenges of use of toxicogenomic data in risk assessment.

7. CONCLUSIONS

This chapter describes the approach that was refined based on performing the DBP case study, summary conclusions of the DBP case study, recommendations, future considerations, and research needs for applying genomic data to risk assessment.

7.1. APPROACH FOR EVALUATING TOXICOGENOMIC DATA IN CHEMICAL ASSESSMENTS

To review, there were two goals of this project (see Chapter 2):

- *Develop a systematic approach that allows the risk assessor to utilize the available toxicogenomic data in chemical-specific health risk assessments performed at U.S. EPA; and*

- *Perform a case study to illustrate the approach.*

The first goal was to develop an approach for evaluating toxicogenomic data in future chemical assessments. The DBP case study was unlike the process for a new risk assessment in a number of ways. In the case study, we had the benefit of utilizing toxicity and human study data set evaluations summarized in the IRIS DBP assessment external review draft. Additionally, the information about DBP from the published literature and the IRIS assessment draft allowed us to focus on one set of endpoints, the male reproductive developmental endpoints. Thus, the case study approach (see Figure 3-1) needed to be refined to develop a systematic approach for incorporating toxicogenomic data in a future chemical assessment (Figure 7-1).

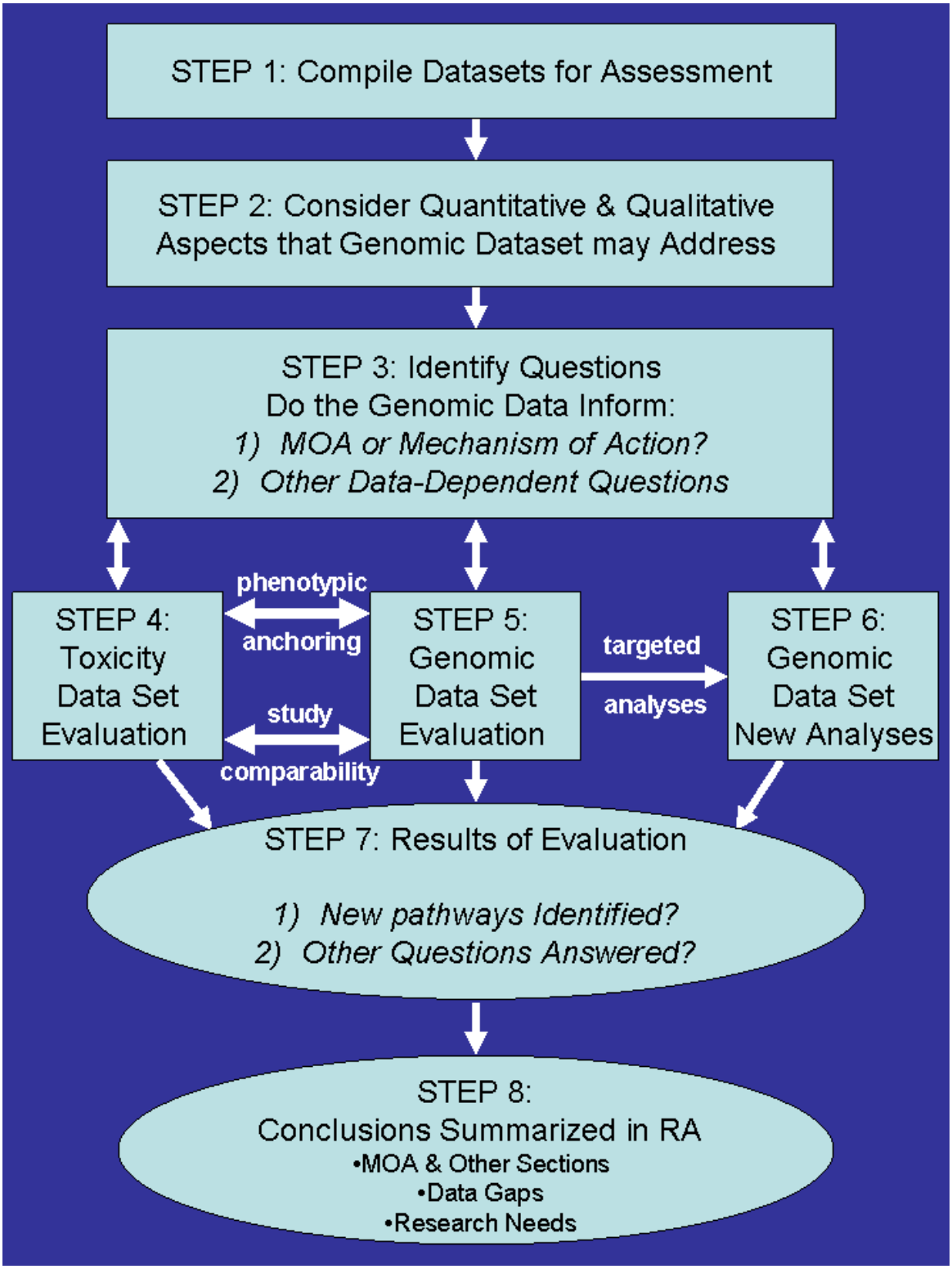

Figure 7-1. Approach for evaluating and incorporating genomic data for health assessments. "Toxicity Data Set Evaluation" may include evaluation of animal toxicity data and/or human outcome data, depending on the available data for the chemical.

The steps of the approach are:

- *STEP 1: Compile the epidemiologic, animal toxicology, and toxicogenomic study data sets.*

- *STEP 2: Consider the quantitative and qualitative aspects of the risk assessment that these data may address.*

  The genomic data set is considered for whether these data could inform risk assessment components (e.g., dose response) and information (e.g., MOA information, interspecies TK differences) useful to risk assessment. The type of information that these data will provide to a risk assessment depends in part on the type of genomic studies (e.g., species, organ, design, method) that are available. A thorough and systematic consideration of the types of information in light of the available genomic data will identify the potential utility of the genomic data and whether these data can be used quantitatively or qualitatively. See Section 3.2 for more details.

- *STEP 3: Formulate questions to direct the toxicogenomic data set evaluation.*

  Questions are formulated that can direct the genomic data evaluation. Some examples of questions considered in the DBP case study are: Do the data inform the MOAs for the female reproductive outcomes?; Do the data inform dose-response? For example, if microarray data are available, then one of the questions will likely include whether the genomic data can inform the mechanism and/or MOA for the chemical as microarray data typically inform the mechanism of action of a chemical. The DBP case study describes some examples and considerations for determining the risk assessment components that may be informed by a particular genomic data set (See Section 3.3 for more details of the considerations).

- *STEPS 4 and 5: Evaluate the toxicity and/or human study and genomic data sets*

  The approach includes an integrated assessment of the toxicogenomic and toxicity data set to relate the affected endpoints (identified in the toxicity data set evaluation) to the pathways (identified in the toxicogenomic data set evaluation) as a method for:

  (1) Determining the level of support for phenotypic anchoring of genomic changes to in vivo outcomes.

  (2) Informing the mechanism of action/MOA.

  Risk assessors may want to utilize aspects of the approach defined herein along with the Mode of Action Framework in the U.S. EPA Cancer Guidelines (U.S. EPA, 2005) and/or other risk assessment decision-logic frameworks for establishing MOAs.

  Another principle of the approach is identifying comparable toxicity and toxicogenomic data. For example, in the DBP case study, all of the toxicogenomic studies were performed in the rat, and, in most cases, the testis. Therefore, the genomic data set was

compared with the rat toxicity data, and focused on effects in the testis. Broadening beyond the DBP example, the available toxicogenomic data are best considered in light of the toxicity or epidemiologic study data with the similarities to the toxicogenomic study design. For example, if toxicogenomic data from human tissue or cells are available, these data are best considered with the human epidemiologic outcome data for the chemical. However, even in the absence of comparable data in the same species, the genomic data may still be utilized, but with less confidence. See Chapters 4 and 5 for further details of the DBP case study toxicity and toxicogenomic data set evaluations.

Chapter 5 includes a number of simple methods for assessing the consistency of the toxicogenomic data. Venn diagrams have been utilized for illustrating the similarities and differences of DEG findings across genomic studies. Figure 5-2 is an example of another method for assessing the consistency of findings across all types of gene expression data.

- *STEP 6: Perform new analyses of the genomic data.*

New analyses of raw toxicogenomic data may be valuable for the assessment depending on the questions asked and the nature of the analyses presented in the published studies. Depending on the pathway-analysis methods used in the published genomic studies, reanalysis with different pathway analysis methods may be warranted. New analyses of the raw data may not be needed—for instance, in the case that the available published data have been analyzed appropriately for application to the specific risk assessment questions. See Chapter 6 for more details of the DBP case study new analyses.

- *STEP 7 and 8: Describe results of evaluations and analyses. Then, summarize these conclusions in the assessment.*

7.2. DBP CASE STUDY FINDINGS

The second goal of the project was to develop a case study. The case study findings are summarized here. The details of the case study evaluation and analyses are presented in Chapters 4−6 (with supplemental material in Appendices A and B). Two advantages to using DBP as the case study chemical are as follows:

- The temporal aspects (e.g., time of dosing and time of evaluation) could be considered because a number of well designed studies exist;
- The expression of a number of the steroidogenesis pathway genes have a strong phenotypic anchoring/association with a number of the male reproductive developmental effects;

- Two well established MOAs for DBP have been defined at the molecular level. DBP is known to affect multiple MOAs allowing for a query of the genomic data for possible MOAs for the unexplained endpoints.

7.2.1. Case Study Question 1: Do the DBP Genomic Data Inform Mechanism of Action and MOA?

In our case study, we found that toxicogenomic data did inform the mechanism of action and MOA. The available genomic and other gene expression data, hormone measurement data, and toxicity data for DBP were instrumental in establishing two of its MOAs: (1) a decrease in fetal testicular T, and (2) a decrease in *Insl3* expression. A decrease in fetal testicular T is the MOA responsible for a number of the male reproductive developmental effects in the rat, and the genomic and other gene expression data identified changes in genes involved in steroidogenesis and cholesterol transport, which is consistent with and provides the underlying basis for the observed decrease in fetal testicular T. A decrease in *Insl3* expression is one of the two MOAs responsible for undescended testis descent, and this MOA is well established by RT-PCR and in vivo toxicology data. RT-PCR studies identified reduced *Insl3* expression (Wilson et al., 2004) after in utero DBP exposure that was associated gubernacular agenesis or abnormalities observed in toxicology studies, effects that are not seen after exposure to chemicals that affect T synthesis or activity (e.g., AR binding). These results provided support for the *Insl3* MOA for DBP.

Rodent reproductive developmental toxicity studies were evaluated for low incidence and low-dose findings as well as for male reproductive development effects that currently do not have a known MOA (see Chapter 4). The testes outcomes were the focus of the case study because the DBP toxicogenomic studies were all performed on testicular tissue. Five testes effects associated with DBP exposure that do not have well described MOAs were identified in this evaluation.

The toxicogenomic and other gene expression studies, including nine published RT-PCR and microarray studies in the rat after in utero DBP exposure (Shultz et al., 2001; Barlow et al., 2003; Lehmann et al., 2004; Wilson et al., 2004; Bowman et al., 2005; Thompson et al., 2004; Thompson et al., 2005; Liu et al., 2005; Plummer et al., 2007), were evaluated. The review of the toxicogenomic data set focused on an evaluation of the consistency of findings from the

published studies, and on whether additional pathways may illuminate the unexplained endpoints. This evaluation found that the gene-level findings from the DBP genomic studies (i.e., microarray, RT-PCR, and protein expression) were highly consistent in both the identification of DEGs and their direction of effect.

New analyses of the Liu et al. (2005) microarray study were performed. These evaluations (see Chapter 5) indicate that there are a number of pathways affected after in utero DBP exposure; some of these pathways are related to new MOAs because they are not related to either the reduced fetal testicular T or the *Insl3* signaling MOAs. The Liu et al. (2005) DBP raw data set was re-analyzed using two different methods, the SNR and the weighted-linear model, both using a statistical cutoff of $p < 0.05$. Each method identified the steroidogenesis and cholesterol transport pathways, thus, corroborating prior study conclusions. Each analysis also identified putative new pathways and processes that are not associated with either *Insl3* or steroidogenesis pathways; some were similar across analytical methods and some were different. The pathways identified that were in common between the two methods (Table 6-3) fall into eight processes (characterized by Ingenuity®): cell signaling, growth and differentiation, metabolism, transcription, immune response, cell adhesion, hormones, and disease. There were 54 pathways, not related to reduced T or *Insl3* expression, including a subset (e.g., WNT signaling and cytoskeleton remodeling) that were not previously identified in the published literature for DBP. One or more of these additional pathways may provide information about the MOAs for the unexplained toxicity endpoints in the rat testes, but this remains to be determined. Evaluating the genomic and toxicity data sets together provided information on potential, heretofore unexplored, MOAs.

There are many possible reasons for the differences in findings between the reanalysis and the published analysis of the Liu et al. (2005) data. These include but are not limited to

(1) The analyses had different purposes. Liu et al. (2005) was interested in determining whether there is a developmental phthalate genomic signature. This work was interested in identifying all affected pathways;

(2) In the 3 years since the study was published, gene and pathway annotation has increased. Further, repeated identification of DEGs and pathways provides an additional level of confidence regarding the importance of "in common" DEGs and pathways but by

no means indicate a lack of importance for the genes and pathways that were not repeatedly identified.

We also asked whether there are appropriate data to develop a regulatory network model for DBP. Using the raw data from Thompson et al. (2005), the only time-course study available at the time of the project, changes in gene expression and pathways were modeled (Figure B-3). Two limitations of these data are that (1) the exposure interval was at the tail end of the critical window of exposure, GD 18, a time that most consider too late to induce the full spectrum of male reproductive developmental effects; and (2) the duration of exposure and developmental time were not aligned because all animals were sacrificed on GD 19 (i.e., the 1 hour time point was the *latest* in development; see Chapter 6 for more discussion). The more recent study of Plummer et al. (2007) may be more appropriate data to use to build a regulatory network model as both time-course of exposure over the critical window of development and microdissection of the testis cell types were employed in their study. Use of these data would allow for a regulatory network model to incorporate both temporal and spatial aspects of DBP's effects on pathways and endpoints.

7.2.2. Case Study Question 2: Do the DBP Genomic Data Inform Interspecies Differences in the TD part of the MOA?

Human gene expression data are not available for DBP. Therefore, the case study used information on interspecies similarities of the affected pathways from other data and methods. We explored the interspecies (rat to human) differences in the TD part of the MOA, focusing on the steroidogenesis pathway underlying the decrease in fetal testicular testosterone MOA. The similarities between genes and protein sequences of genes in the biosynthesis of steroid pathway suggest similarities in the pathway across humans and rats. Comparisons of the steroidogenesis genes and pathway were performed to evaluate cross-species similarity metrics (see Chapter 6) using three approaches: (1) protein sequence similarity; (2) pathway network similarities; and (3) promoter-region conservation. Results from all three approaches indicate that steroidogenesis pathways are relatively highly conserved across rats and humans and, thus, qualitatively, the rat and human mechanisms for steroidogenesis share many similarities.

These results further corroborate what is known about the similar roles for androgens during normal male development in both rat and human. However, the data sources used for all three approaches have gaps in the knowledge bases. For the pathway network diagramming, there is a data quality concern. Due to data quality caveats, it is difficult to use these new lines of evidence to quantitatively inform the relative sensitivity to DBP across species. It is possible that the small differences across species have a strong penetrance, leading to significant differences in what proteins may be more sensitive to DBP for T production. Because there are some questions as to the reliability of the data used to generate the pathway comparisons used for each species, there is no basis on which to transform a measure of conservation to a quantitative measure of sensitivity. Thus, we do not recommend utilizing these data to inform interspecies uncertainty in the case of DBP because it is difficult to make unequivocal conclusions regarding a "high" versus "low" degree of conservation for the genes in this pathway based on these data alone. These methods, however, when based on high quality data, could be applied quantitatively to future chemical assessments.

We further considered whether some steroidogenesis genes are of higher relative importance and, thus, should be weighted higher in a cross-species assessment of the steroidogenesis pathway. The initiating event for DBP action in the male reproductive developmental outcomes has not been established. Some knowledge of the rate-limiting steps for steroidogenesis, in the unperturbed scenario, is available. P450scc has been identified in some studies as a limiting enzymatic step for T production (Miller, 1988; Omura and Morohashi, 1995). However, the information on kinetics reflects the unperturbed state because the rate-limiting step was defined in assays without DBP exposure. Additionally, the rate-limiting step information is limited in scope to steroidogenic enzymes and not all upstream activities leading to T production, such as STAR, a protein that impacts the availability of cholesterol (by transporting cholesterol to the inner mitochondrial membrane for cleavage by P450scc) for T production. Thus, there is no a priori knowledge to argue for placing more weight on a particular gene leading to T production.

While the confidence in the cross species comparisons of the steroidogenesis pathway were not high enough to utilize the findings quantitatively, the findings do add to the weight-of-evidence suggesting that the role of T in male fetal development in rats and humans is well conserved. Further, the exploratory methods for developing metrics for cross-species pathway

similarities described in this document (see Chapter 6) may be developed and validated in the future for quantitative use in risk assessment.

7.2.3. Application of Genomic Data to Risk Assessment: New Methods

None of the DBP genomic studies were designed with the application to risk assessment in mind. Microarray and other 'omic data analytical methods were originally developed for screening purposes (i.e., designed to err on the side of false positives over false negatives). For risk-assessment application, different genomic analytical tools are needed that do not err on the side of false positives (i.e., do not detecting a change in gene expression by chance) and reliably separate signal from noise. In traditional pathway level analysis, significant genes are mapped to their respective pathways. Depending on whether the number of genes that map to any given pathway, the role of the pathway can be over of underestimated. To overcome this problem, we developed the overall pathway activity (OPA) method that employs one as opposed to two steps (i.e., first, identifying DEGs and second, identifying significantly affected pathways by grouping the DEGs using pathway analysis programs). This method, that ranks pathways based on the expression level of all genes in a given pathway, shows promise for use in risk assessment but needs to be further validated.

Chapter 6 describes exploratory methods for developing a genetic regulatory network model and measuring cross-species differences for a given pathway. Genetic regulatory network models can be very useful for understanding the temporal sequence of critical biological events perturbed after chemical exposure, and thus, useful to a risk assessment. We developed a method for developing a genetic regulatory network model for DBP based on the available data. The availability of a time-course data (Thompson et al. [2005]) enabled our group to model the series of events that occurred between exposure to DBP and the onset of toxic reproductive outcomes by the generation of a regulatory network model. However, given the limitations of the Thompson et al. (2005) study design, we did not draw conclusions about affected genes and pathways over time for DBP from this study. Given the limitations of the Thompson et al. (2005) data (see Chapter 6), the exercise allowed us to develop methods for analyzing time course data for use in building a regulatory network model. We used three different methods to assess rat-to-human conservation as metrics that may inform the interspecies differences for one

MOA, the reduced fetal testicular T. However, there are a number of challenges in using similarity scores to quantitatively estimate the human relevance of a MOA (section 6.5.).

7.2.4. Application of Genomic Data to Risk Assessment: Using Data Quantitatively

This case study was limited to qualitative uses of genomics in risk assessment. U.S. EPA and the larger scientific community working with genomics are interested in methods to use genomic data quantitatively in risk assessment. Genomic data were not assessed quantitatively in this case study due to the absence of dose-response global gene expression studies (i.e., microarray studies) for DBP. There is one dose-response RT-PCR study that, although not a genomic (i.e., not global) study, was considered for use quantitatively in risk assessment (Lehmann et al., 2004; Table 7-1). Strengths of the Lehmann et al. (2004) study include the following:

Table 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

	0.1 mg/kg-d	1 mg/kg-d	10 mg/kg-d	30 mg/kg-d	50 mg/kg-d	80 mg/kg-d	100 mg/kg-d
Precursor event[a]	↓ *Hsd3b*	↓ *Hsd3b* ↓*Scarb1*	NC in gene exp. NC in [T]	ND for gene exp. NC in [T]	↓*Scarb1* ↓*Hsd3b* ↓*StAR* ↓*Cyp11a1* ↓[T]	ND for gene exp.	↓*Scarb1* ↓*3β-HSD* ↓*StAR* ↓*P450scc* ↓[T]
in vivo endpoint						↑ incidence of absent, poorly developed, or atrophic testis and underdeveloped or absent epididymis[b]	Retained nipples and areolae[c]

NC, no statistically significant change; ND = not determined (Lehmann et al. (2004) did not test 80 mg/kg-d).

Sources: [a]Lehmann et al. (2004); [b]NTP, 1991; [c]Mylchreest et al., 2000.

- The study includes low to high doses.

- Some of the genes assessed in this study were first identified in microarray studies, providing a level of connection between the gene and particular outcomes as well as demonstrating reproducibility across studies. For example, findings for *Star* gene expression are reproduced across protein expression, RT-PCR, and microarray studies.

However, there are a number of issues in utilizing these dose-response RT-PCR data. These limitations include the following:

- Some of the gene expression changes are not reproducible. For example, *Kit* was observed to be significantly altered in the Lehmann et al. (2004) study but was not observed to be significantly reduced after in utero DBP exposure in a microarray study (Liu et al., 2005) utilizing the Affymetrix® gene chip, yet *Kit* is on the Affymetrix® rat chip.

- The relationship between statistical significance and biological significance is not known for genomic data. For example, the expression of *Hsd3b* mRNA is statistically significantly altered at lower doses than a statistically significant [T] decrease was observed. Thus, Lehmann et al. (2004) argued that the changes in *Hsd3b* at 0.1 and 1.0 mg/kg-d were not biologically significant. It is also not known whether changes in the expression of a single or multiple steroidogenesis genes would lead to a significant alteration in [T] and the phenotype.

- Inter-litter variability could not be characterized from the Lehmann et al. (2004) data because the RT-PCR data were collected on five individual pups representing four to five litters per treatment group (i.e., ~1 pup/litter). In order to have appropriate data for BMD modeling, litter mean values calculated from a study with a greater sample size and multiple litters are needed to allow characterization of inter-litter variability.

Regarding quantitative measures of intraspecies and interspecies differences, it should be noted that the same information which is necessary for quantitative assessment of interspecies differences (Section 7.2.2) may be useful for characterizing intraspecies variability, and vice versa. In particular, factors that explain or predict interstrain differences in rodent sensitivity to DBP, such as those noted between Wistar and SD rats, may be hypothesized to contribute to human variability. Further, toxicologically important interstrain differences identified from the toxicogenomic data could be an excellent data source for investigating whether they are also important for modulating interspecies sensitivity.

7.3. LESSONS LEARNED

The lessons learned from the case study are grouped by research needs and recommendations that are useful to research scientists and those who work on risk assessments.

7.3.1. Research Needs

7.3.1.1. *Data Gaps and Research Needs: DBP*

There are some research needs that would be very useful to a DBP risk assessment. Research needs for DBP include the following:

(1) *Developing a genetic regulatory network model using the Plummer et al. (2007) data.* This data set would be an excellent source of temporal and spatial gene expression information because one of its studies includes three time intervals, thus covering the entire critical window for male reproductive outcomes; and a second study used microdissection of the cord and interstitial cells of the testis. This study was not modeled because it was not published until after the modeling work was performed. By comparing gene expression, they hypothesized the MOA underlying the gonocyte and LC effects. These data could be used to develop a regulatory network for DBP in utero exposure and effects on the rat testis;

2) *Performing microarray studies in male reproductive tissues, other than the testis, affected by DBP in order to understand the similarities and differences in DBP-affected pathways in across reproductive organs and tissues in the male rat.* Bowman et al. (2005) performed such a study in the WDs, but studies in other male tissues are needed;

3) *Performing microarray studies in human tissues (either cell lines or from aborted male fetal tissue), along with parallel in vitro and in vivo studies in rats for validation and comparison.* Such data would provide critical information for the IRIS DBP assessment on qualitative, and possibly quantitative, interspecies differences in TDs sensitivity. Some human studies found an association between in utero phthalate exposure and newborn male reproductive developmental measures (Swan et al., 2005; Main et al., 2006) that indicate human relevance for some of the DBP effects observed in male rat studies;

4) *Performing well designed proteomic and metabolomic studies to understand the affect of in utero DBP exposure on the function of expressed proteins, and on cellular metabolites.* These data may provide complementary data to the available transcriptomic data, which could yield some new insights;

5) *Performing genomic studies to identify early, critical, upstream events as a means to identify the initiating event for DBP's action in the testis.* This would require performing studies much earlier in gestation, at the beginning of sexual differentiation. In addition,

such studies may require greater sensitivity regarding gene expression change identification because a statistically significant change may be greater than a biologically significant change. If identified, the initiating event could be utilized in the risk assessment, thereby reducing uncertainty;

6) *Performing genomic studies to understand whether the female reproductive tract malformations after DBP exposure share a common MOA with the male development reproductive effects.* This line of research would identify pathways affected in the developing female reproductive tracts after early gestational DBP exposure.

7) *Comparing the affected DEGs and pathways between the phthalates with and without developmental effects could be useful for a cumulative risk assessment of the developmental phthalates.* All of the data from the Liu et al. (2005) data set could be utilized to evaluate this issue. Further, evaluating consistency of findings across chemicals in the same MOA class that do and do not produce the same set of effects could be useful for improving specificity of the MOA findings.

7.3.1.2. *Research Needs for Toxicity and Toxicogenomic Studies for Use in Risk Assessment: Future Chemical Assessments*

The U.S. EPA and the larger scientific community are interested in methods to use genomic data quantitatively in risk assessment. This case study was limited to qualitative uses of genomics in risk assessment due to the absence of dose-response global gene expression studies (i.e., microarray studies) for DBP. Thus, multiple dose microarray studies are needed (Table 7-2). Such studies are very costly and without proper design and power can be difficult to interpret because the lower doses may not affect gene expression in every organ assessed, leading to the need for increased sample size. For example, 500 mg/kg-d DBP was used as the single dose in the published microarray studies because exposure during the critical window at this dose leads to the maximum reproductive developmental effects (i.e., almost all animals are affected in every male pup) without effects on maternal toxicity. In a dose-response study including low to high doses, the sample size per dose group would need to be high enough to increase statistical power (i.e., the detection of gene expression changes when only a few animals are affected). For example, if an endpoint is affected in 20% of the animals at lower doses, then the sample size for microarray studies must be large enough to identify the affected animals (with affected gene expression). Perhaps the highest priority study is one that assesses global gene expression and toxicity endpoints of interest; the testis would be collected at GD 19 in one group of animals but a second group would be followed through to evaluation of the developmental endpoint of interest.

Table 7-2 describes some of the priority research needs for toxicogenomic studies for developmental toxic chemicals, including DBP. First, appropriate time-course gene-expression data over the critical window, using a small subset of genes whose altered expression is linked to the outcome of interest, would be very relevant for developing a regulatory network model.

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

Purpose	Study Needed
1) Develop a regulatory network model	Exposure time-course microarray data.
2) Improve pathway analysis statistical power	Number of replicates increased.
3) Use of toxicogenomic data to inform toxicokinetics in dose-response analysis	Genomic and toxicity studies with same study design: Generate TK data in relevant study (time, dose, tissue), and obtain relevant internal dose measure to derive best internal dose metric.
4) Use of toxicogenomic data in dose-response analysis	Multiple doses in microarray studies in parallel with phenotypic anchoring.
5) Phenotypic anchoring; informing MOA (Figure 3-4)	Similar study design characteristics for genomic and toxicity studies (i.e., dose, timing of exposure, organ/tissue evaluated).
6) Assess intraspecies differences	A study assessing multiple doses across rat strains (e.g., Wistar vs. SD); endpoint and microarray component of the study.
7) Assess interspecies differences	A study to assess whether different species with similar pathways (genes and sequence of steps) have a similar sensitivity to a given chemical. The findings could potentially enhance the utility of TgX data to aid species extrapolation in risk assessments.
8) Appropriate statistical pathway analysis methods for use in risk assessment	Further comparisons and evaluations of different methods.
9) Screening and categorizing chemicals by MOA in risk assessment (e.g., cumulative risk assessment)	Genomic (transcriptomic, proteomic, and/or metabolomic) signatures can be particularly useful for screening and categorizing chemicals by MOA in risk assessment.

These studies need to be carefully designed based on the information on the critical window of exposure and the relationship to the particular outcome of concern. Second, the statistical power of pathway-analysis methods for global expression techniques, including microarrays, proteomics and metabolomics, could be improved by designing and performing studies with more replicates. Thus, variability would be better characterized. Third, it would be helpful to design genomic studies that could inform both TKs and dose response (#3 and #4, Table 7-2).

Performing genomic and toxicity studies with similar designs would provide useful information. These studies would be designed at the most relevant time of exposure, include low to high doses, and assess the relevant tissues. Relevant internal dose measurements could be obtained on which to base the internal dose metric. These studies, employing genomic and toxicity studies of comparable designs, would allow for phenotypic anchoring of dose, gene expression, and outcome, and thus, could potentially be used in dose-response analysis. Studies with both a toxicity and toxicogenomic component would obviously require assessment of a large sample size to be informative. These same studies could be used to inform MOA (#5) and could be adapted to comparing species (#6). Finally, further development and comparison studies to identify appropriate statistical pathway analysis methods for use in risk assessment are needed (#8). It is important to note that such studies require research funding and laboratories with expertise in both genomics and toxicology.

Research needs for toxicity studies that would improve the utility in risk assessment are described in Table 7-3. As was noted for the DBP case (Chapter 4), complete reporting is necessary for studies that are intended for use in risk assessment.

7.3.2. Recommendations

Based on the lessons learned from performing the DBP case study exercise, we developed some recommendations or best practices for performing assessments for chemicals having available genomic data. We recommend following the principles of the approach described herein, to thoroughly consider the available genomic data for whether it can inform every information type useful to risk assessment, and to evaluate genomic

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

Study Aspect	Research Need
Study design	Exposing animals during optimal developmental stage/time (i.e., for the critical window).
	Assessing outcome at optimum developmental stage/time for that outcome.
	Parallel study design characteristics with toxicogenomic studies (i.e., dose, timing of exposure, organ/tissue evaluated) to obtain comparable toxicity and toxicogenomic studies to aid connections between gene expression changes and outcomes.
Reporting	Individual animal data to aid identification of low incidence effects, correlate gene expression changes and outcomes, and characterize intraspecies variability.
	All endpoints that were evaluated (independent of whether the outcome was positive or negative).

data and toxicity data together to assess phenotypic anchoring. In addition, we recommend four specific methods for evaluating genomic data that arose from the DBP case study. Two of these recommendations are straightforward and could reasonably be performed by a risk assessor with basic genomics training:

1) *Evaluate the genomic and other gene expression data for consistency of findings across studies to provide a weight-of-the-evidence (WOE) evaluation of the affected gene expression and pathways.* Some simple methods, such as using Venn diagrams and gene-expression compilation approaches can be applied to risk assessment. When evaluating the consistency of toxicogenomic data findings, it was advantageous to include all of the available gene expression data (single gene, global gene expression, protein, RNA) because the single gene expression techniques have been traditionally used to confirm the results of global gene expression studies.

2) *Perform benchmark dose (BMD) modeling on high-quality RT-PCR dose-response studies for genes known to be in the causal pathway of a MOA or outcome of interest.* Obtaining a BMD and BMDL (benchmark dose lower confidence limit) is a useful starting point for both linear low-dose extrapolation and reference value approaches. We are not indicating which approach is appropriate to take for making predictions about the potential risk below the BMD or BMDL. "High quality" is defined in this context as a well conducted study that assessed enough animals and litters for sufficient statistical power for characterizing the mean responses and the variability (interlitter and intralitter variability).

Two additional recommendations require expertise in genomic data analysis methods to implement:

3) *Perform new analysis of toxicogenomic raw data in order to identify all affected pathways or for other risk assessment applications.* Most often, microarray studies are conducted for different purposes (e.g., basic science, pharmaceutical development). In these cases, new pathway analysis of microarray data can be potentially useful.

4) *Develop a genetic regulatory network model for the chemical of interest to define the system of interacting regulatory DNA sequences, expression of genes, and pathways for one or more outcomes of interest.* Genetic regulatory network model methods, developed as part of this case study, could be used in a risk assessment. If time-course genomic data are available, the temporal sequence of mechanistic events after chemical exposure can be defined, and the earliest affected genes and pathways, that may be define the initiating event, may be identified.

7.3.3. Application of Genomic Data to Risk Assessment: Future Considerations

A number of the issues that emerged in evaluating the DBP genomic data set are relevant to using genomic data in risk assessment in general. Some issues regarding the use of genomic data are to the same as for the use of precursor information in risk assessment, regardless of the technique used to gather the information. Two outstanding questions are

- *How is the biologically significant level of change in a precursor marker determined?* And, specifically for toxicogenomic data, *what are the key genes (i.e., a key gene, a handful of genes associated with the outcome of interest, a genomic signature) whose altered expression leads to an adverse outcome?* Currently, decisions about the degree of change of a precursor event tend to be based on statistical significance because data to address biological significance are typically lacking (as is the case for T levels and male development of the testis). Genes are identified as DEGs in microarray studies based on statistical-significance criteria that may not reflect biological significant changes (i.e., identified genes may not be biologically meaningful while unidentified genes may be

meaningful). This point is also relevant to the question: *What pathway analysis methods are most appropriate for risk assessment?* As noted in Chapter 6, it is difficult to know whether one has identified the biologically relevant DEGs and pathways. Statistically significant changes and repeated findings of the same genes and pathways across studies and using different analytical methods does not necessarily provide a greater confidence regarding biological significance of these genes and pathways over other genes and pathways. Further, there is a bias towards the well annotated genes as biologically significant when, in fact, the unannotated genes could be of greater importance.

- *What are the requirements for linkage of precursor events to in vivo endpoints?* Studies to assess the relationship between the gene expression and outcomes are needed to establish a causal connection.

There are also a number of technical issues in utilizing microarray data in U.S. EPA risk assessments that have not fully been surmounted. The primary technical issue is the validation of the reproducibility of microarray study results. Reproducibility depends on biological sample preparation, interlaboratory (presumably related to operator and protocol differences), intralaboratory (presumably related to operator differences), and platform variability. The results of the MAQC project (see Chapters 2 and 5) revealed that reproducibility was achieved when using the same biological sample. This is very encouraging for using microarray data in risk assessment. However, biological sample variability still needs to be addressed in order that protocols and details of the underlying reasons for the variability can be understood.

A number of the issues stem from the complexity of the data output from the global expression techniques (e.g., microarrays, proteomics, metabolomics). This is in part a training issue. To address the training needs, the U.S. EPA Risk Assessment Forum held introductory and intermediate level training in genomics in 2007. The FDA has also held genomics training (http://www.fda.gov/cder/genomics/Default.htm). However, it would be advantageous for U.S. EPA to embark on further training of risk assessors to enable them to perform analyses of microarray and other genomic data analysis techniques,and to understand the issues in applying traditional analytical methods to risk assessment.

If additional case studies are performed using the approach outlined in Figure 7-1, we recommend a chemical whose exposure leads to both cancer and noncancer outcomes to explore use of these data for multiple outcomes as well as the impacts on the different risk assessment

paradigms and processes (e.g., cancer versus noncancer). In fact, one of the phthalates might be a good candidate chemical for such a case study. Further, performing case studies on data-rich and data-poor chemicals would aid in further evaluating the approach described herein.

The approach for utilizing toxicogenomic data in risk assessment outlined in this document may be applied to other chemical assessments. This document advances the effort to devise strategies for using genomic data in risk assessment by defining an approach, performing a case study, and defining critical issues that need to be addressed to better utilize these data in risk assessment. This case study serves as an example of the considerations and methods for using genomic data in future risk assessments for environmental agents.

8. REFERENCES

Anderson, NL; Anderson, NG. (1998) Proteome and proteomics: new technologies, new concepts, and new words. Electrophoresis 19(11):1,853–1,861.

Aardema, MJ; MacGregor, JT. (2002) Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies. Mutat Res 499:13–25.

Andersen, ME; Thomas, RS; Gaido, KW; et al. (2005) Dose-response modeling in reproductive toxicology in the systems biology era. Reprod Tox 19(3):327–337.

Anonymous. (2006) Making the most of microarrays (editorial). Nature Biotech 24(9): 1039.doi: 10.1038/nbt0906-1039.

Ball, CA; Brazma, A; Causton, H; et al. (2004) Submission of microarray data to public repositories. PLoS Biol 2(9):E317.

Barlow, NJ; Foster, PM. (2003) Pathogenesis of male reproductive tract lesions from gestation through adulthood following in utero exposure to di(n-butyl) phthalate. Toxicol Pathol 31(4):397–410.

Barlow, NJ; Phillips, SL; Wallace, DG; et al. (2003) Quantitative changes in gene expression in fetal rat testes following exposure to di(n-butyl) phthalate. Toxicol Sci 73(2):431–441.

Barlow, NJ; McIntyre, BS; Foster, PM. (2004) Male reproductive tract lesions at 6, 12, and 18 months of age following in utero exposure to di(n-butyl) phthalate. Toxicol Pathol 32:79–90.

BASF. (1992) Study on the oral toxicity of dibutyl phthalate in Wistar rats. Administration via the diet over 3 months. Prepared by BASF Corporation, Ludwigshafen, Germany for Eastman Kodak Company; Project No. 31S0449//89020; Submitted under TSCA Section 8D; EPA Document No. 86-920000903; NTIS OTS0535640 (As cited in NTP-CERHR, 2000).

Bodin, L; Beaune, PH; Loriot, MA. (2005) Determination of cytochrome P450 2D6 (CYP2D6) gene copy number by real-time quantitative PCR. J Biomed Biotechnol 2005(3):248–253.

Boverhof, DR; Zacharewski, TR. (2006) Toxicogenomics in risk assessment: applications and needs. Toxicol Sci 89(2):352–360.

Bowman, CJ; Turner, KJ; Sar, M; et al. (2005) Altered gene expression during rat Wolffian duct development following di(n-butyl) phthalate exposure. Toxicol Sci 86:161–174.

Brazma, A; Hingamp, P; Quackenbush, J; et al. (2001) Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. Nature Genetics 29(4):365–371.

Buckley, PG; Mantripragada, KK; Piotrowski, A; et al. (2005) Copy-number polymorphisms: mining the tip of an iceberg. Trends Genet 21(6):315–317.

Canales, RD; Luo, Y; Willey, JC; et al. (2006) Evaluation of DNA microarray results with quantitative gene expression platforms. Nat Biotechnol 24(9):1,115–1,122. doi:10.1038/nbt1236. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt1236.pdf.

Carruthers, CM; Foster, P. (2005) Critical window of male reproductive tract development in rats following gestational exposure to di-n-butyl phthalate. Birth Defects Res B Dev Reprod Toxicol 74(3):277–285.

Casanova, M; Morgan, KT; Gross, EA; et al. (1994) DNA-protein cross-links and cell replication at specific sites in the nose of F344 rats exposed subchronically to formaldehyde. Fundam Appl Toxicol 23(4):525−536.

Casciano, DA; Woodcock, J. (2006) Empowering microarrays in the regulatory setting (forward). Nat Biotechnol 24(9):1,103.

Ceder, R; Merne, M; Staab, C; et al. (2007) The application of normal, SV40 T-antigen-immortalised and tumour-derived oral keratinocytes, under serum-free conditions, to the study of the probability of cancer progression as a result of environmental exposure to chemicals. Altern Lab Anim 35(6):621−639.

Chan, VS; Theilade, MD. (2005) The use of toxicogenomic data in risk assessment: a regulatory perspective. Clin Toxicol 43(2):121–126.

Coen, M; Lenz, EM; Nicholson, JK; et al. (2003) An integrated metabonomic investigation of acetaminophen toxicity in the mouse using NMR spectroscopy. Chem Res Toxicol 16(3):295–303.

Cunningham, ML; Bogdanffy, MS; Zacharewski, T; et al. (2003) Workshop overview: use of genomic data in risk assessment. Toxicol Sci 73(2):209–215.

Dalgaard, M; Nellemann, C; Lam, HR; et al. (2001) The acute effects of mono(2-ethylhexyl)phthalate (MEHP) on testes of prepubertal Wistar rats. Toxicol Lett 122(1):69–79.

Daston, GP. (2007) Genomics and developmental risk assessment. Birth Defects Res A Clin Mol Teratol 79:1–7.

David, RM. (2006) Proposed mode of action for in utero effects of some phthalate esters on the developing male reproductive tract. Toxicol Pathol 34(3):209–19. Available online at http://tpx.sagepub.com/cgi/reprint/34/3/209.pdf.

Delbes, G; Levacher, C; Duquenne, C; et al. (2005) Endogenous estrogens inhibit mouse fetal Leydig cell development via estrogen receptor alpha. Endocrinol 146(5):2454−2461.

Demoulin, JB; Ericsson, J; Kallin, A; et al. (2004) Platelet-derived growth factor stimulates membrane lipid synthesis through activation of phosphatidylinositol 3-kinase and sterol regulatory element-binding proteins. J Biol Chem 279(34):35,392−35,402.

Dix, DJ; Gallagher, K; Benson, WH; et al. (2006) A framework for the use of genomics data at the EPA. Nat Biotechnol 24(9):1,108–1,111. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt0906-1108.pdf.

Donato, MT; Lahoz, A; Castell, JV; et al. (2008) Cell lines: a tool for in vitro drug metabolism studies. Curr Drug Metab 9(1):1−11.

Ekins, S; Nikolsky, Y; Nikolskaya, T. (2005) Techniques: application of systems biology to absorption, distribution, metabolism, excretion and toxicity. Trends Pharm Sci 26(4):202–209.

Ellinger-Ziegelbauer, H; Stuart, B; Wahle, B; et al. (2005) Comparison of the expression profiles induced by genotoxic and nongenotoxic carcinogens in rat liver. Mutat Res 575(1–2):61–84.

El-Masri, H; Constan, AA; Ramsdell, HS; et al. (1996) Physiologically based pharmacodynamic modeling of an interaction threshold between trichloroethylene and 1,1-dichloroethylene in Fischer 344 rats. Toxicol Appl Pharmacol 141:124–132.

El-Masri, H; Bell, DA; Portier, CJ. (1999) Effects of glutathione transferase theta polymorphism on the risk estimates of dichloromethane to humans. Toxicol Appl Pharmacol 158(3):221–230.

Ema, M; Amano, H; Itami, T; et al. (1993) Teratogenic evaluation of di-n-butyl phthalate in rats. Toxicol Lett 69(2):197−203.

Ema, M; Amano, H; Ogawa, Y. (1994) Characterization of the developmental toxicity of di-n-butyl phthalate in rats. Toxicology 86(3):163–174.

Ema, M; Kurosaka, R; Amano, H; et al. (1995) Developmental toxicity evaluation of mono-n-butyl phthalate in rats. Toxicol Lett 78(2):101−106.

Ema, M; Kurosaka, R; Harazono, A; et al. (1996). Phase specificity of developmental toxicity after oral administration of mono-n-butyl phthalate in rats. Arch Environ Contam Toxicol 31(2):170−176.

Ema, M; Miyawaki, E; Kawashima, K. (1998) Further evaluation of developmental toxicity of di-n-butyl phthalate following administration during later pregnancy in rats. Toxicol Lett 98(1–2):87–93.

Ema, M; Miyawaki, E. (2001a) Adverse effects on development of the reproductive system in male offspring of rats given monobutyl phthalate, a metabolite of dibutyl phthalate, during late pregnancy. Reprod Toxicol 15(2):189−194.

Ema, M; Miyawaki, E. (2001b) Effects of monobutyl phthalate on reproductive function in pregnant and pseudopregnant rats. Reprod Toxicol 15(3):261−267.

Ema, M; Harazono, A; Miyawaki, E; et al. (1997) Developmental effects of di-n-butyl phthalate after a single administration in rats. J Appl Toxicol 17(4):223−229.

Ema, M; Miyawaki, E; Kawashima, K. (2000a) Effects of dibutyl phthalate on reproductive function in pregnant and pseudopregnant rats. Reprod Toxicol. 14(1):13−19.

Ema, M; Miyawaki, E; Kawashima, K. (2000b) Critical period for adverse effects on development of reproductive system in male offspring of rats given di-n-butyl phthalate during later pregnancy. Toxicol Lett 111(3):271–278.

Fang, H; Tong, W; Perkins, R; et al. (2005) Bioinformatics approaches for cross-species liver cancer analysis based on microarray gene expression profiling. BMC Bioinformatics 6 (Suppl 2):S6.

FDA (Food and Drug Administration). (2003) Guidance for industry. Pharmacogenomic data submissions. Draft guidance. Washington, DC: Food and Drug Administration. Available online at www.fda.gov/cder/guidance/5900dft.pdf [accessed 26 July 2004].

Fennell, TR; Krol, WL; Sumner, SC; et al. (2004) Pharmacokinetics of dibutylphthalate in pregnant rats. Toxicol Sci 82(2):407−418.

Ferrara, D; Hallmark, N; Scott, H; et al. (2006) Acute and long-term effects of in utero exposure of rats to di(n-butyl) phthalate on testicular germ cell development and proliferation. Endocrinol. 147(11):5352–5362.

Fisher, JS; Macpherson, S; Marchetti, N; et al. (2003) Human 'testicular dysgenesis syndrome': a possible model using in-utero exposure of the rat to dibutyl phthalate. Human Reprod 18(7):1389–1394.

Forst, CV; Schulten, K. (1999) Evolution of metabolisms: a new method for the comparison of metabolic pathways using genomics information. J Comput Biol 6(3−4):343−360.

Forst, CV; Schulten, K. (2001) Phylogenetic analysis of metabolic pathways. J Mol Evol 52(6):471−489.

Foster, PMD. (2005) Mode of action: impaired fetal Leydig cell function—effects on male reproductive development produced by certain phthalate esters. Crit Rev Toxicol 35(8–9):713–719.

Frueh, FW; Huang, SM; Lesko, LJ. (2004) Regulatory acceptance of toxicogenomics data. Environ Health Perspect 112(12):A663–A664.

Frueh, FW. (2006) Impact of microarray data quality on genomic data submissions to the FDA. Nature Biotechnology 24(9):1,105–1,107. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt0906-1105.pdf.

Geng, J; Strobel, HW. (1995) Identification of inducible mixed function oxidase system in rat glioma C6 cell line. J Neurochem 65(2):554−563.

Ginsberg, G; Smolenski, S; Hattis, D; et al. (2002) Population distribution of aldehyde dehydrogenase-2 genetic polymorphism: implications for risk assessment. Regul Toxicol Pharmacol 36(3):297–309.

Girotti, M; Zingg, HH. (2003) Gene expression profiling of rat uterus at different stages of parturition. Endocrinol. 144(6):2,254−2,265.

Golub, TR; Slonim, DK; Tamayo, P; et al. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286(5439):531−537.

Gray,.LE, Jr; Ostby, JS; Kelce, WR. (1994) Developmental effects of an environmental antiandrogen: the fungicide vinclozolin alters sex differentiation of the male rat. Toxicol Appl Pharmacol 129:46–52.

Gray, LE, Jr; Ostby, J; Monosson, E; et al. (1999a) Environmental antiandrogens: low doses of the fungicide vinclozolin alter sexual differentiation of the male rat. Toxicol Indust Health 15(1–2):48–64.

Gray, LE, Jr; Wolf, C; Lambright, C; et al. (1999b) Administration of potentially antiandrogenic pesticides (procymidone, linuron, chlozolinate, *p,p*'-DDE, and ketoconazole) and toxic substances (dibutyl- and diethylhexyl phthalate, PCB 169, and ethane dimethane sulphonate) during sexual differentiation produces diverse profiles of reproductive malformations in the male rat. Toxicol Indust Health 15(1–2):94–118.

Gray, LE, Jr; Ostby, J; Furr, J; et al. (2000) Perinatal exposure to the phthalates DEHP, BBP, and DINP, but not DEP, DMP, or DOTP, alters sexual differentiation of the male rat. Toxicol Sci 85(2):350−365.

Gray, LE, Jr.; Ostby, J; Furr, J; et al. (2001) Effects of environmental antiandrogens on reproductive development in experimental animals. Hum Reprod Update. 7(3):248−264.

Guerreiro, N; Staedtler, F; Grenet, O; et al. (2003) Toxicogenomics in drug development. Toxicol Pathol 31(5):471−479.

Guo, L; Lobenhofer, EK; Wang, C; et al. (2006) Rat toxicogenomic study reveals analytical consistency across microarray platforms. Nat Biotechnol 24(9):1162–1169. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt1238.pdf

Hackett, JL; Lesko, LJ. (2003) Microarray data—the US FDA, industry, and academia. Nat Biotechnol 21(7):742–743.

Heymans, M; Singh, AK. (2003) Deriving phylogenetic trees from the similarity analysis of metabolic pathways. Bioinformatics 19(Suppl 1):i138-146.

Hong, SH; Kim, TY; Lee, SY. (2004) Phylogenetic analysis based on genome-scale metabolic pathway reaction content. Appl Microbiol Biotechnol 65(2):203–210.

Hamadeh, HK; Bushel, PR; Jayadev, S; et al. (2002a) Gene expression analysis reveals chemical-specific profiles. Toxicol Sci 67(2):219–231.

Hamadeh, HK; Bushel, PR; Jayadev, S; et al. (2002b) Prediction of compound signature using high density gene expression profiling. Toxicol Sci 67(2):232–240.

Hedberg, JJ; Grafstrom, RC; Vondracek, M; et al. (2001) Micro-array chip analysis of carbonyl-metabolizing enzymes in normal, immortalized and malignant human oral keratinocytes. Cell Mol Life Sci 2001;58(11):1719−1726.

Hellwig, J; van Ravenzwaay, M; Mayer, M; et al. (2000) Pre- and postnatal oral toxicity of vinclozolin in Wistar and Long-Evans rats. Regul Toxicol Pharmacol 32:42–50.

Higuchi, TT; Palmer, JS; Gray, LE, Jr; et al. (2003) Effects of dibutyl phthalate in male rabbits following in utero, adolescent, or postpubertal exposure. Toxicol Sci 72(2):301−313.

Ji, H; Davis, RW. (2006) Data quality in genomics and microarrays. Nat Biotechnol 24(9):1112–1113. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt0906-1112.pdf.

Kakuma, T; Lee, Y; Higa, M; et al. (2002) Leptin, troglitazone, and the expression of sterol regulatory element binding proteins in liver and pancreatic islets. Proc Natl Acad Sci USA 97(15):8,536−8,541.

Kavlock, R; Ankley, GT; Collette, T; et al. (2005) Computational Toxicology: Framework, Partnerships, and Program Development: September 29–30, 2003, Research Triangle Park, North Carolina. Reprod Toxicol 19(3):265-280.

Kurhekar, MP; Adak, S; Jhunjhunwala, S; et al. (2002) Genome-wide pathway analysis and visualization using gene expression data. Pac Symp Biocomput 462−473.

Kelce, WR; Monosson, E; Gamcsik, MP; et al. (1994) Environmental hormone disruptors: evidence that vinclozolin developmental toxicity is mediated by antiandrogenic metabolites. Toxicol Appl Pharmacol 126(2):276−285.

Kennedy, S. (2002) The role of proteomics in toxicology: identification of biomarkers of toxicity by protein expression analysis. Biomarkers 7(4):269–290.

Kim, T; Shin, J; Lee, S; et al. (2004) Effects of in utero exposure of diethylstilbestrol and dibutyl phthalate on the testis descent in rat offspring. Abstract # 574. Toxicologist 78:118.

Klaunig, J; Babich, M; Baetcke, K; et al. (2003) PPARalpha agonist-induced rodent tumors: modes of action and human relevance. Crit Rev Toxicol 33(6):655–780.

Kleymenova, E; Dei Rossi, E; Liu, K; et al. (2004) Developmental exposure to di-n-butyl phthalate affects cord organization and Sertoli cell-gonocyte interactions in the fetal rat testis. Abstract # 926. Toxicologist 78:190.

Kleymenova, E; Swanson, C; Gaido, KW. (2005a) Low dose cellular responses in the fetal rat testis in utero exposed to di-n-butyl phthalate. Abstract #1372. Toxicologist 79:281.

Kleymenova, E; Swanson, C; Boekelheide, K; et al. (2005b) Exposure in utero to di(n-butyl) phthalate alters the vimentin cytoskeleton of fetal rat Sertoli cells and disrupts Sertoli cell-gonocyte contact. Biol Reprod 73(3):482–490.

Klonisch, T; Fowler, PA; Hombach-Klonisch, S. (2004) Molecular and genetic regulation of testis descent and external genitalia development. Dev Biol 270:1–18.

Lambright, C; Ostby, J; Bobseine, K; et al. (2000) Cellular and molecular mechanisms of action of linuron: an antiandrogenic herbicide that produces reproductive malformations in male rats. Toxicol Sci 56(2):389–399.

Lee, K-Y; Shibutani, M; Takagi, H; et al. (2004) Diverse developmental toxicity of di-n-butyl phthalate in both sexes of rat offspring after maternal exposure during the period from late gestation through lactation. Toxicology 203(1–3):221–238.

Lehmann, KP, Phillips, S; Sar, M; et al. (2004) Dose-dependent alterations in gene expression and testosterone synthesis in the fetal testes of male rats exposed to di (n-butyl) phthalate. Toxicol Sci 81(1):60–68.

Leighton, JK. (2005) Application of emerging technologies in toxicology and safety assessment: regulatory perspectives. Int J Toxicol 24(3):153–155.

Li, AC; Glass, CK. (2004) PPAR- and LXR-dependent pathways controlling lipid metabolism and the development of atherosclerosis. J Lipid Res 45(12):2,161−2,173.

Liang, CP; Tall, AR. (2001) Transcriptional profiling reveals global defects in energy metabolism, lipoprotein, and bile acid synthesis and transport with reversal by leptin treatment in ob/ob mouse liver. J Biol Chem 276(52):49,066−49,076.

Liang, S; Fuhrmann, S; Somogyi, R. (1998) Reveal, a general reverse engineering algorithm for inference of genetic network architectures. Pac Symp Biocomput 18−29.

Liu, K; Lehmann, KP; Sar, M; et al. (2005) Gene expression profiling following in utero exposure to phthalate esters reveals new gene targets in the etiology of testicular dysgenesis. Biol Reprod 73(1):180–192.

Lobenhofer, EK; Bushel, PR; Afshari, CA; et al. (2001) Progress in the application of DNA microarrays. Environ Health Perspect 109(9):881–891.

Mahood, IK; Hallmark, N; McKinnell, C; et al. (2005) Abnormal Leydig cell aggregation in the fetal testis of rats exposed to di (n-butyl) phthalate and its possible role in testicular dysgenesis. Endocrinol 146(2):613–623.

Main, KM; Mortensen, GK; Kaleva, MM; et al. (2006) Human breast milk contamination with phthalates and alterations of endogenous reproductive hormones in infants three months of age. Environ Health Perspect. 114(2):270−276.

MAQC Consortium; Shi, L; Reid, LH; et al. (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24(9):1,151−1,161. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt1239.pdf.

Miller, WL. (1988) Molecular biology of steroid hormone synthesis. Endocr Rev 9(3):295−318.

McKinnell, C; Sharpe, RM; Mahood, K; et al. (2005) Expression of insulin-like factor 3 protein in the rat testis during fetal and postnatal development and in relation to cryptorchidism induced by in utero exposure to di (n-Butyl) phthalate. Endocrinology 146(10):4,536−4,544.

Moggs, JG; Tinwell, H; Spurway, T; et al. (2004) Phenotypic anchoring of gene expression changes during estrogen-induced uterine growth. Environ Health Perspect 112(16):1,589−1,606.

Moggs, JG. (2005) Molecular responses to xenoestrogens: mechanistic insights from toxicogenomics. Toxicology 213(3):177–193.

Mylchreest, E; Cattley, RC; Foster, PMD. (1998) Male reproductive tract malformations in rats following gestational and lactational exposure to di(n-butyl) phthalate: an antiandrogenic mechanism? Toxicol Sci 43:47–60.

Mylchreest, E; Sar, M; Cattley, RD; et al. (1999) Disruption of androgen-regulated male reproductive development by di(n-butyl) phthalate during late gestation in rats is different from flutamide. Toxicol Appl Pharmacol 156(2):81–95.

Mylchreest, E; Wallace, DG; Cattley, RC; et al. (2000) Dose-dependent alterations in androgen-regulated male reproductive development in rats exposed to di(n-butyl) phthalate during late gestation. Toxicol Sci 55:143–151.

Mylchreest, E; Sar, M; Wallace, DG; et al. (2002) Fetal testosterone insufficiency and abnormal proliferation of Leydig cells and gonocytes in rats exposed to di(n-butyl) phthalate. Repro Toxicol 16:19–28.

Naciff, JM; Hess, KA; Overmann, GJ; et al. (2005) Gene expression changes induced in the testis by transplacental exposure to high and low doses of 17α-ethynyl estradiol, genistein, or bisphenol A. Toxicol Sci 86(2):396–416.

Naciff, JM; Jump, ML; Torontali SM; et al. (2002) Gene expression profile induced by 17alpha-ethynyl estradiol, bisphenol A, and genistein in the developing female reproductive system of the rat. Toxicol Sci 68(1):184–199.

Naciff, JM; Overmann, GJ; Torontali, SM; et al. (2003) Gene expression profile induced by 17 alpha-ethynyl estradiol in the prepubertal female reproductive system of the rat. Toxicol Sci 72(2):314–330.

NRC (National Research Council) (2007a) Applications of toxicogenomic technologies to predictive toxicology and risk assessment. Washington, DC: National Academy Press.

NRC (National Research Council). (2007b) Toxicity testing in the 21st century. Washington, DC: National Academy Press.

NTP (National Toxicology Program). (1991) Final report on the reproductive toxicity of di-n-butyl phthalate (Cas. No. 84-74-3) in Sprague-Dawley rats. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Toxicity Program. Research Triangle Park, NC. NTP Report # RACB89012. NTIS # PB92-111996.

NTP (National Toxicology Program). (1995) NTP technical report on toxicity studies of dibutyl phthalate administered in feed to F344/N rats and B6C3F1 mice. Toxicity Report Series Number 30. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Toxicity Program. Research Triangle Park, NC. Available online at http://ntp.niehs.nih.gov/ntp/htdocs/ST_rpts/tox030.pdf.

NTP-CERHR (National Toxicology Program – Center for the Evaluation of Risks to Human Reproduction). (2000) NTP-CERHR monograph on the potential human reproductive and developmental effects of di-n-butyl phthalate (DBP). National Toxicology Program, U.S. Department of Health and Human Services, Research Triangle Park, NC. Available online at http://cerhr.niehs.nih.gov/chemicals/phthalates/dbp/DBP_Monograph_Final.pdf

NTP-CERHR (National Toxicology Program – Center for the Evaluation of Risks to Human Reproduction). (2003) Monograph on the potential human reproductive and developmental effects of di-n-butyl phthalate (DBP). NTP Cerhr Mon (4):i-III90.

Oberemm, A; Onyon, L; Gundert-Remy, U. (2005) How can toxicogenomics inform risk assessment? Toxicol Appl Pharmacol 207(2 Suppl):592–598.

Omura, T; Morohashi, K. (1995) Gene regulation of steroidogenesis. J Steroid Biochem Mol Biol 53(1-6):19−25.

Ostby, J; Kelce, WR; Lambright, C; et al. (1999) The fungicide procymidone alters sexual differentiation in the male rat by acting as an androgen-receptor antagonist in vivo and in vitro. Toxicol Indust Health 15(1–2):80–93.

Pandey, A; Mann, M. (2000) Proteomics to study genes and genomes. Nature 405(6,788):837–846.

Parks, LG; Ostby, JS; Lambright, CR; et al. (2000) The plasticizer diethylhexyl phthalate induces malformations by decreasing fetal testosterone synthesis during sexual differentiation in the male rat. Toxicol Sci 58(2):339−349.

Pavlidis, P; Lewis, DP; Noble, WS. (2002) Exploring gene expression data with class scores. Pac Symp Biocomput 474−485.

Patterson, TA; Lobenhofer, EK; Fulmer-Smentek, FB; et al. (2006) Performance comparison of one-color and two-color platforms within the Microarray Quality Control (MAQC) project. Nat Biotechnol 24(9):1,140–1,150. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt1242.pdf

Pennie, WD; Pettit, SD; Lord, PG. (2004) Toxicogenomics in risk assessment – An overview of an ILSI HESI collaborative research program. Environ Health Perspect 112(4):417–419.

Pettit, SD; Lord, PG; Pennie, WD. (2003) The application of genomics to mechanism-based risk assessment—overview of an ILSI HESI collaborative research program. J Toxicol Sci 28(4):220–221.

Plummer, S; Hallmark, N; Sharpe, R; et al. (2005) Identification of gene clusters and signaling pathways affected by dibutyl phthalate—nuclear receptor interactions in fetal rat testes. Abstract #2258. Toxicologist 79:463.

Plummer, S; Sharpe, RM; Hallmark, N; et al. (2007) Time-dependent and compartment-specific effects of in utero exposure to di(n-butyl) phthalate on gene/protein expression in the fetal rat testis as revealed by transcription profiling and laser capture microdissection. Toxicol Sci 97(2):520–352.

Raunio, H; Hakkola, J; Hukkanen, J; et al. (1999) Expression of xenobiotic-metabolizing CYPs in human pulmonary tissue. Exp Toxicol Pathol 51(4−5):412−417.

Reynolds, VL. (2005) Applications of emerging technologies in toxicology and safety assessment. Int J Toxicol 24(3):135–137.

Robinson, DE; Pettit, SD; Morgan, DG. (2003) Use of genomics in mechanism based risk assessment. In: Inoue, T; Pennie, WD; eds., Toxicogenomics. Tokyo: Springer-Verlag, pp. 194–203.

Saillenfait, AM; Payan, JP; Fabry, JP; et al. (1998) Assessment of the developmental toxicity, metabolism, and placental transfer of Di-n-butyl phthalate administered to pregnant rats. Toxicol Sci 45(2):212−224.

Schlosser, PM; Lilly, PD; Conolly, RB; et al. (2003) Benchmark dose risk assessment for formaldehyde using airflow modeling and a single-compartment, DNA-protein cross-link dosimetry model to estimate human equivalent doses. Risk Anal 23(3):473−487.

Shamir, R; Maron-Katz, A; Tanay, A; et al. (2005) EXPANDER--an integrative program suite for microarray data analysis. BMC Bioinformatics 6:232.

Shea-Eaton, WK; Trinidad, MJ; Lopez, D; et al. (2001) Sterol regulatory element binding protein-1a regulation of the steroidogenic acute regulatory protein gene. Endocrinol 142(4):1,525-1,533.

Shastry, BS. (2006) Pharmacogenetics and the concept of individualized medicine. Pharmacogenomics J 6:16–21.

Shi, L; Tong, W; Goodsaid, F; et al. (2004) QA/QC: challenges and pitfalls facing the microarray community and regulatory agency. Expert Rev Mol Diagn 4(6):761–777.

Shippy, R; Fulmer-Smentek, S; Jensen, RV; et al. (2006). Using RNA sample titrations to assess microarray platform performance and normalization techniques. Nat Biotechnol 24(9):1123–1131. Available online at http://www.nature.com/nbt/journal/v24/n9/abs/nbt1241.html.

Shultz, VD; Phillips, S; Sar, M; et al. (2001) Altered gene profiles in fetal rat testes after in utero exposure to di(n-butyl) phthalate. Toxicol Sci 64(2):233–242.

Sladek, R; Bader, JA; Giguere, V. (1997) The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. Mol Cell Biol 17(9): 5,400−5,409.

Simmons, PT; Portier, CJ. (2002) Toxicogenomics: the new frontier in risk analysis. Carcinogenesis 23(6):903–905.

Smith, LL. (2001) Key challenges for toxicologists in the 21st century. Trends Pharmacol Sci 22(6):281–285.

Staab, CA; Ceder, R; Roberg, K; et al. (2008) Serum-responsive expression of carbonyl-metabolizing enzymes in normal and transformed human buccal keratinocytes. Cell Mol Life Sci 65(22):3,653−3,663.

Stocco, DM. (2002) Clinical disorders associated with abnormal cholesterol transport: mutations in the steroidogenic acute regulatory protein. Mol Cell Endocrinol 191(1):19–25.

Swan, SH; Main, KM; Liu, F; et al. (2005) Decrease in anogenital distance among male infants with prenatal phthalate exposure. Environ Health Perspect 113(8):1,056–1,061. Erratum in: Environ Health Perspect 2005 Sep 113(9):A583.

Swanson, HI. (2004) Cytochrome P450 expression in human keratinocytes: an aryl hydrocarbon receptor perspective. Chem Biol Interact 149(2-3):69−79.

Thomas, RS; Allen, BC; Nong, A; et al. (2007) A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. Toxicol Sci 98(1):240−248.

Thompson, CJ; Ross, SM; Gaido, KW. (2004) Di(n-butyl) phthalate impairs cholesterol transport and steroidogenesis in the fetal rat testis through a rapid and reversible mechanism. Endocrinology 145(3):1,227–1,237.

Thompson, CJ; Ross, SM; Hensley, J; et al. (2005) Differential steroidogenic gene expression in the fetal adrenal versus the testis and rapid and dynamic response of the fetal testis to di(n-butyl) phthalate. Biol Reprod 73(5):908–917.

Tilton, SC; Orner, GA; Benninghoff, AD; et al. (2008) Genomic profiling reveals an alternate mechanism for hepatic tumor promotion by perfluorooctanoic acid in rainbow trout. Environ Health Perspect 116(8):1,047–1,055. Available online at http://www.ehponline.org/docs/2008/11190/abstract.html

Tomfohr, J; Lu, J; Kepler, TB. (2005) Pathway level analysis of gene expression using singular value decomposition. BMC Bioinformatics 6:225.

Tong, W; Lucas, AB; Shippy, R; et al. (2006) Evaluation of external RNA controls for the assessment of microarray performance. Nat Biotechnol 24(9):1132–1139. Available online at http://www.nature.com/nbt/journal/v24/n9/pdf/nbt1237.pdf.

Tuma, DJ; Casey, CA. (2003) Dangerous byproducts of alcohol breakdown--focus on adducts. Alcohol Res Health. 27(4):285−290.

Turner, KJ; McIntyre, BS; Phillips, SL; et al. (2003) Altered gene expression during rat Wolffian duct development in response to in utero exposure to the antiandrogen linuron. Toxicol Sci 74(1):114–128.

Turner, ME; Martin, C; Martins, AL; et al. (2007) Genomic and expression analysis of multiple Sry loci from a single Rattus norvegicus Y chromosome. BMC Genet 8:11.

U.S. EPA (Environmental Protection Agency). (2002) Interim genomics policy. Prepared by the Science Policy Council, Washington, DC. Available online at http://www.epa.gov/osa/spc/genomics.htm.

U.S. EPA (Environmental Protection Agency). (2004a) A framework for a computational toxicology research program. Office of Research and Development, National Center for Computational Toxicology, Washington, DC. Available online at http://www.epa.gov/comptox/comptox_framework.html.

U.S. EPA (Environmental Protection Agency). (2004b) Potential implications of genomics for regulatory and risk assessment applications at EPA. Genomics Task Force Workgroup, Science Policy Council, Washington, DC; EPA 100/B-04/002. Available online at http://www.epa.gov/OSA/pdfs/EPA-Genomics-White-Paper.pdf.

U.S. EPA (Environmental Protection Agency). (2004c) Acetochlor report of the Cancer Assessment Review Committee (CARC). Fourth Evaluation. Office of Pesticide Programs, Washington, DC.

U.S. EPA (Environmental Protection Agency). (2005) Supplemental guidance for assessing susceptibility from early-life exposure to carcinogens. Risk Assessment Forum, Washington, DC, EPA/630/R-03/003F. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=160003.

U.S. EPA (Environmental Protection Agency). (2006a) Toxicological review of dibutyl phthalate (di-n-butyl phthalate). External peer review draft. Integrated Risk Information System, Washington, DC; NCEA-S-1755.

U.S. EPA (Environmental Protection Agency). (2006b) Interim guidance for microarray-based assays: data submission, quality, analysis, management and training considerations. External review draft. Genomics Tasks Force Workgroup, Science Policy Council, Washington, DC. Available online at http://www.epa.gov/osa/spc/genomicsguidance.htm.

U.S. EPA (Environmental Protection Agency). (2006c) Summary of the NCEA colloquium on current use and future needs of genomics in ecological and human health risk assessment (final report). National Center for Environmental Assessment, Washington, DC; EPA/600/R-04/039. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=149984.

U.S. EPA (Environmental Protection Agency). (2006d) Revised reregistration eligibility decision document (RED) for MSMA; DSMA, and CAMA and cacodylic acid. Office of Prevention, Pesticides and Toxic Substances, Washington, DC; EPA/738/R-06/021. Available online at http://www.epa.gov/oppsrrd1/REDs/organic_arsenicals_red.pdf.

U.S. EPA (Environmental Protection Agency). (2006e) Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment. National Center for Environmental Assessment, Washington, DC; EPA/600/R-05/043F. Available online at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=157668.

van Pelt, AM; de Rooij, DG; van der Burg, B; et al. (1999) Ontogeny of estrogen receptor-beta expression in rat testis. Endocrinology 140(1):478−483.

Vinggaard, AM; Nellemann, C; Dalgaard, M; et al. (2002) Antiandrogenic effects in vitro and in vivo of the fungicide prochloraz. Toxicol Sci 69(2):344–353.

Vondracek, M; Weaver, DA; Sarang, Z; et al. (2002) Transcript profiling of enzymes involved in detoxification of xenobiotics and reactive oxygen in human normal and simian virus 40 T antigen-immortalized oral keratinocytes. Int J Cancer 99(6):776−782.

Vondracek, M; Xi, Z; Larsson, P; et al. (2001) Cytochrome P450 expression and related metabolism in human buccal mucosa. Carcinogenesis 22(3):481−488.

Waters, MD; Fostel, JM. (2004) Toxicogenomics and systems toxicology: aims and prospects. Nat Rev Genet 5(12):936–948. Available online at http://www.cmbi.ru.nl/edu/bioinf4/articles/pdf/tox_ReviewToxicogenomicsNaturedec2004.pdf.

Wilson, VS; Wood, CR; Held, GA; et al. (2003) Comparison of the effects of two AR antagonists on androgen dependent tissue weight and hormone levels in male rats and on expression of three androgen dependent genes in the ventral prostate. Toxicologist 72:131.

Wilson, VS; Lambright, C; Furr, J; et al. (2004) Phthalate ester-induced gubernacular lesions are associated with reduced insl3 gene expression in the fetal rat testis. Toxicol Lett 146(3):207–215.

Wine, RN; Li, L-H; Barnes, LH; et al. (1997). Reproductive toxicity of di-*n*-butylphthalate in a continuous breeding protocol in Sprague-Dawley rats. Environ Health Perspect 105:102–107.

Wong, C; Kelce, WR; Sar, M; et al. (1995) Androgen receptor antagonist versus agonist activities of the fungicide vinclozolin relative to hydroxyflutamide. J Biol Chem 270(34):19998–20003.

Wu, W; Hu, W; Kavanagh, JJ. (2002) Proteomics in cancer research. Int J Gynecol Cancer 12(5):409–423.

Wyde, ME; Kirwan, SE; Zhang, F; et al. (2005) Di-n-butyl phthalate activates constitutive androstane receptor and pregnane X receptor and enhances the expression of steroid-metabolizing enzymes in the liver of rat fetuses. Toxicol Sci 86(2):281–290.

Xiong, S; Chirala, SS; Wakil, SJ. (2000) Sterol regulation of human fatty acid synthase promoter I requires nuclear factor-Y- and Sp-1-binding sites. Proc Natl Acad Sci USA 97(8):3948−3953.

Yu, X; Griffith, WC; Hanspers, K; et al. (2006) A system-based approach to interpret dose- and time-dependent microarray data: quantitative integration of gene ontology analysis for risk assessment. Toxicol Sci 92(2):560–577.

Zhang, Y; Jiang, X; Chen, B. (2004) Reproductive and developmental toxicity in F1 Sprague-Dawley male rats exposed to di-n-butyl phthalate in utero and during lactation and determination of its NOAEL. Reprod Toxicol 18(5):669–676.

Zien, A; Kuffner, R; Zimmer, R; et al. (2000) Analysis of gene expression data with pathway scores. Proc Int Conf Intell Syst Mol Biol 8:407−417.

URLS AND WEBSITES:

http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=149984

http://cerhr.niehs.nih.gov/news/phthalates/DEHP-final.pdf

http://cerhr.niehs.nih.gov/news/phthalates/report.html

http://david.abcc.ncifcrf.gov/list.jsp

http://intranet.epa.gov/ncea/pdfs/qmp/ncea_qmp.pdf

http://searchlauncher.bcm.tmc.edu/help/BLASToutput.html#anchor14684156.

http://www.ehponline.org/txg/docs/admin/txg-n-press.html?section=toxicogenomics

http://www.genome.jp/kegg

http://www.hesiglobal.org/Committees/ TechnicalCommittees/Genomics/EBI+Toxicogenomics.htm

http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=114215&ordinalpos=3&itool=EntrezSystem2.PEntrez.Gene.Gene_ResultsPanel.Gene_RVDocSum

http://www.ncbi.nlm.nih.gov/sites/entrez?db=PubMed

www.epa.gov/iris/whatsnewarchive.htm

www.genego.com

www.omics.org

9. GLOSSARY

Amplified Fragment Length Polymorphism PCR (AFLP-PCR or AFLP): A PCR-based tool used in genetics research, DNA fingerprinting, and in the practice of genetic engineering.

Benchmark Dose (BMD) or Concentration (BMC): A dose or concentration that produces a predetermined change in response rate of an adverse effect (called the benchmark response or BMR) compared to background.

Copy Number Polymorphism (CNP): Normal variation in the number of copies of a sequence within the DNA.

Complementary DNA (cDNA): A double stranded DNA version of an mRNA molecule.

Exposure: Contact made between a chemical, physical, or biological agent and the outer boundary of an organism. Exposure is quantified as the amount of an agent available at the exchange boundaries of the organism (e.g., skin, lungs, gut).

Exposure Assessment: An identification and evaluation of the human population exposed to a toxic agent, describing its composition and size, as well as the type, magnitude, frequency, route and duration of exposure.

Expressed Sequence Tag (EST): A short subsequence of a transcribed cDNA sequence.

Gene Ontology (GO): A collaborative project of the Gene Ontology Consortium that has developed three structured controlled vocabularies (ontologies) that describe gene products in terms of their associated biological processes, cellular components and molecular functions in a species-independent manner. There are three separate aspects to this effort: first, the development and maintenance of the ontologies themselves; second, the annotation of gene products, which entails making associations between the ontologies and the genes and gene products in the collaborating databases; and third, development of tools that facilitate the creation, maintenance and use of ontologies.

Gene Regulatory Network (GRN) Model: A representation of the regulation (e.g., positive or negative regulation) of genes and their expression (e.g., RNAs, proteins, metabolites) of a system (e.g., cell, tissue), and their relations. A GRN model can be expressed at the genomic and metabolic level. Genes can be viewed as nodes in the network, with input being proteins (e.g., transcription factors), and outputs being the level of gene expression. Further, GRNs can describe changes over time or space if based on time course or spatial compartment data.

Genomics: The study of the genome and include genome sequencing and genotype analysis techniques (e.g., polymorphism identification).

Hazard Assessment: The process of determining whether exposure to an agent can cause an increase in the incidence of a particular adverse health effect (e.g., cancer, birth defect) and whether the adverse health effect is likely to occur in humans.

Hazard Characterization: A description of the potential adverse health effects attributable to a specific environmental agent, the mechanisms by which agents exert their toxic effects, and the associated dose, route, duration, and timing of exposure.

Key Event: An empirically observable precursor step that is, itself, a necessary element of the mode of action or is a biologically based marker for such an element (U.S. EPA, 2005).

Lowest Observed Adverse Effect Level (LOAEL): The lowest exposure level at which there are biologically significant increases in frequency or severity of adverse effects between the exposed population and its appropriate control group.

Lowest Observed Effect Level (LOEL): In a study, the lowest dose or exposure level at which a statistically or biologically significant effect is observed in the exposed population compared with an appropriate unexposed control group.

Microarray Quality Control (MAQC): A project that was developed to provide quality-control tools to the microarray community in order to avoid procedural failures and to develop guidelines for microarray data analysis by providing the public with large reference data sets along with readily accessible reference RNA samples.

Metabolomics: Metabolomics is the study of low-molecular-weight metabolic products.

Microarray: A microarray is a tool for analyzing gene expression that consists of a small membrane or glass slide containing samples of many genes arranged in a regular pattern.

Mechanism of Action: The complete molecular sequence of events between the interaction of the chemical with the target site and observation of the outcome. Thus, the mechanism of action can include toxicokinetic and/or toxicodynamic steps.

Mode of Action (MOA): One event, or a sequence of key events, that the outcome is dependent upon (i.e., part of the causal pathway and not a coincident event).

No Observed Adverse Effect Level (NOAEL): The highest exposure level at which there are no biologically significant increases in the frequency or severity of adverse effect between the exposed population and its appropriate control; some effects may be produced at this level, but they are not considered adverse or precursors of adverse effects.

No Observed Effect Level (NOEL): An exposure level at which there are no statistically or biologically significant increases in the frequency or severity of any effect between the exposed population and its appropriate control.

Omics: Omics is a general term for a broad discipline of science and engineering for analyzing the interactions of biological information objects in various ‘omes’ such as toxicogenome, proteome, and metabolome.

Physiologically Based Pharmacokinetic (PBPK) Model: A model that estimates the dose to a target tissue or organ by taking into account the rate of absorption into the body, distribution among target organs and tissues, metabolism, and excretion.

Principal Component Analysis (PCA): A technique for analysis of multivariate data that is similar to SVD (see below). There is a direct relation between PCA and SVD in the case where principal components are calculated from the covariance matrix. Compared to PCA, SVD is more fundamental because SVD simultaneously provides the PCAs in both row and column spaces.

Proteomics: The study of proteins in an organism.

Reverse Transcription Polymerase Chain Reaction (RT-PCR): A two-step process for converting RNA to DNA and the subsequent PCR amplification of the reversely transcribed DNA.

Human Health Risk Assessment: The evaluation of scientific information on the hazardous properties of environmental agents (hazard characterization), the dose-response relationship (dose-response assessment), and the extent of human exposure to those agents (exposure assessment). The product of the risk assessment is a statement regarding the probability that populations or individuals so exposed will be harmed and to what degree (risk characterization).

Serial Analysis of Gene Expression (SAGE): A powerful tool that allows the analysis of overall gene expression patterns with digital analysis.

Single-Nucleotide Polymorphism (SNP): A DNA sequence variation occurring when a single nucleotide — A, T, C, or G — in the genome (or other shared sequence) differs between members of a species (or between paired chromosomes in an individual).

Singular value decomposition (SVD): A technique for analysis of multivariate data. This method describes a system of high number of correlated variables by uncorrelated reduced

number of variables. For analysis of microarray data, SVD provides a linear projection of the gene expression data from the genes × samples space to a noise reduced space and thus, differentiates underlying signals from the noise. Noise reduced space approximates the data with a fraction of the overall expression.

Toxicogenomics: A set of technologies for assessing the genome, transcriptome, proteome, and metabolome gene products after toxic agent exposure.

Transcriptomics: A set of techniques to measure global mRNA expression; it is a tool used to understand specific the expression of genes and pathways involved in biological processes.

APPENDIX A

SUPPORTING TABLES FOR CHAPTER 5

Appendix A contains additional tables that support the work shown in Chapter 5.

Table A-1. Weight of evidence (WOE) for statistically significant gene expression changes after in utero exposure to dibutyl phthalate (DBP) from the whole rat testis microarray studies[a] as reported in Thompson et al. (2005)[b], Shultz et al. (2001)[b], Liu et al. (2005)[c,d], and Plummer et al. (2007)[e]

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Aacs	Acetoacetyl-CoA synthetase	GD 12−19	Down	−0.37 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Aadat	Aminoadipate aminotransferase	GD 12−19	Down	−0.38 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Abcg1	ATP-binding cassette, sub-family G (WHITE), member 1	GD 12−19	Up	0.38 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Acaa1	Acetyl-Coenzyme A acyltransferase 1	GD 12−19	Down	−0.37 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Acaca	Acetyl-Coenzyme A carboxylase alpha	GD 12−19	Down at GD 19	≥2	2-fold	Shultz et al., 2001
Acadl	Acetyl-Coenzyme A dehydrogenase, long-chain	GD 12−19	Down at GD 19	≥2	2-fold	Shultz et al., 2001
Acads	Acyl-Coenzyme A dehydrogenase, short chain	GD 12.5−15.5	Up	1.50	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Acsl4	Acyl-CoA synthetase long-chain family member 4	GD 12−19	Down	−0.60 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Adam15	A disintegrin and metallopeptidase domain 15 (metargidin)	GD 12.5−17.5	Up	1.20	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Adamts1	A disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 1	GD 12.5−19.5	Down	−1.35	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Admr	Adrenomedullin receptor	GD 12−19	Down	−0.90 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Adra1b	Adrenergic receptor, alpha 1b	GD 12−19	Down	−0.30 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Akt2	Thymoma viral proto-oncogene 2	GD 12−21	Down at GD 21	≥2	2-fold	Shultz et al., 2001
Alas1	Aminolevulinic acid synthase 1	GD 12−19	Down	−1.01 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Alas1	Aminolevulinic acid synthase 1	GD 12.5−17.5	Down	−1.33	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Alas1	Aminolevulinic acid synthase 1	GD 12.5−19.5	Down	−1.44	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Aldh1a3	Aldehyde dehydrogenase family 1, subfamily A3	GD 12−19	Down	−0.43 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Aldh2	Aldehyde dehydrogenase 2	GD 12−19	Down	−0.82 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Aldh2	Aldehyde dehydrogenase 2	GD 12.5−17.5	Down	−1.50	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Aldh2	Aldehyde dehydrogenase 2	GD 12.5−19.5	Down	−1.91	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Aldoa	Aldolase A, fructose-bisphosphate	GD 12.5−19.5	Down	−1.24	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Aldoc	Aldolase C	GD 12−19	Down	−0.44 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Anxa5	Annexin A5	GD 12.5−19.5	Down	−1.20	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Aox1	Aldehyde oxidase 1	GD 12−19	Down	−0.50 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Aqp1	Aquaporin 1	GD 12.5−15.5	Down	−1.29	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Arf3	ADP-ribosylation factor 3	GD 12.5−17.5	Down	−1.23	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Arrb2	Arrestin, beta 2	GD 12−21	Down at GD 21	$\geq$2	2-fold	Shultz et al., 2001
Asns	Asparagine synthetase	GD 12−19	Down	−0.24 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ass	Argininosuccinate synthetase	GD 12−19	Down	−0.82 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Atf2	Activating transcription factor 2	GD 12−21	Up at GD 21	$\geq$2	2-fold	Shultz et al., 2001

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Atf4	Activating transcription factor 4	GD 19 for 3 hr	Up after 3 hr	0.67	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Atp1b1	ATPase, Na^{+}/K^{+} transporting, beta 1 polypeptide	GD 12−19	Down	−0.24 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Atp4b	ATPase, H^{+}/K^{+} exchanging, beta polypeptide	GD 12−19	Down	−0.60 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Atp5f1	ATP synthase, H^{+} transporting, mitochondrial F0 complex, subunit B1	GD 12.5−15.5	Up	1.22	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Baiap2	Brain-specific angiogenesis inhibitor 1-associated protein 2	GD 12−19	Down	−0.22 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Bhlhb2	Bhlhb2 basic helix-loop-helix domain containing, class B2	GD 19 for 3 hr	Up after 3 hr	0.88	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Bhmt	Betaine-homocysteine methyltransferase	GD 12−19	Down	−0.24 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Birc5	Baculoviral IAP repeat-containing 5	GD 12.5−15.5	Up	1.68	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Btg2	B-cell translocation gene 2, anti-proliferative	GD 19 for 1 hr	Up after1 hr	1.30	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Btg2	B-cell translocation gene 2, anti-proliferative	GD 19 for 3 hr	Up after 3 hr	1.88	$p < 0.05$ (ANOVA)	Thompson et al., 2005
C4a	Complement component 4a	GD 19 for 6 hr	Down after 6 hr	−0.77	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cadps	Ca^{2+}-dependent secretion activator	GD 12−19	Up	0.31 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Calb2	Calbindin 2	GD 12−19	Down	−0.77 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cd63	CD63 antigen	GD 12.5−19.5	Down	−1.36	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cdkn1c	Cyclin-dependent kinase inhibitor 1C (P57)	GD 12−19	Down	−0.81 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Cdkn1c	Cyclin-dependent kinase inhibitor 1C (P57)	GD 19 for 6 hr	Down after 6 hr	−1.08	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cdkn1c	Cyclin-dependent kinase inhibitor 1C (P57)	GD 18−19 for 18 hr	Down after 18 hr	1.63	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cebpb	CCAAT/enhancer binding protein (C/EBP), beta	GD 12−19	Down	−0.6 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cebpd	CCAAT/enhancer binding protein (C/EBP), delta	GD 19 for 3 hr	Up after 3 hr	1.62	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Clu	Clusterin	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Clu	Clusterin	GD 18 for 18 hr	Up after 18 hr	1.03	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cmklr1	Chemokine-like receptor 1	GD 12.5−19.5	Down	−1.17	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cnr1	Cannabinoid receptor 1 (brain)	GD 19 for 3 hr	Up after 3 hr	0.99	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cnbp	Cellular nucleic acid binding protein	GD 12.5−19.5	Down	−1.29	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cpa1	Carboxypeptidase A1	GD 12.5−17.5	Down	−1.73	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cpa1	Carboxypeptidase A1	GD 12.5−19.5	Down	−2.33	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cpd	Carboxypeptidase D	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Cpe	Carboxypeptidase E	GD 12−19	Up	0.59 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cpt1a	Carnitine palmitoyltransferase 1a, liver	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Cpt1a	Carnitine palmitoyltransferase 1a, liver	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Cpt1b	Cpt1b carnitine palmitoyltrans-ferase1b, muscle	GD 12−19	Up	0.23 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cpz	Carboxypeptidase Z	GD 12−19	Up	0.21 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Crabp2	Cellular retinoic acid binding protein 2	GD 12−19	Down	−0.31 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Crabp2	Cellular retinoic acid binding protein 2	GD 19 for 6 hr	Down after 6 hr	−1.24	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Crem	cAMP responsive element modulator	GD 19 for 3 hr	Up after 3 hr	0.58	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Crispld2	Cysteine-rich secretory protein LCCL domain containing 2	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cryab	Crystallin, alpha B	GD 12−19	Up	0.22 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ctgf	Connective tissue growth factor	GD 19 for 3 hr	Up after 3 hr	2.10	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Ctgf	Connective tissue growth factor	GD 19 for 6 hr	Up after 6 hr	2.37	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Ctsb	Cathepsin B	GD 12.5−15.5	Up	1.53	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Ctsd	Cathepsin D	GD 12.5−19.5	Down	−1.22	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cxcl10	Chemokine (C-X-C motif) ligand 10	GD 19 for 3 hr	Up after 3 hr	2.07	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cyb5	Cytochrome b-5	GD 12−19	Down	−0.30 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	GD 12−19	Down	−1.07 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	GD 18 for 18 hr	Down after 18 hr	−1.93	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	GD 12.5−17.5	Down	−1.71	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	GD 12.5−19.5	Down	−2.85	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cyp11b1	Cytochrome P450, subfamily 11B, polypeptide 1	GD 18 for 18 hr	Down after 18 hr	−1.63	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	GD 12−19	Down	−1.76 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	GD 18 for 18 hr	Down after 18 hr	−2.1	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	GD 12.5−17.5	Down	−2.15	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	GD 12.5−19.5	Down	−3.08	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cyp51	Cytochrome P450, subfamily 51	GD 18 for 18 hr	Down after 18 hr	−1.06	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Cyp51	Cytochrome P450, subfamily 51	GD 12.5−17.5	Down	−1.59	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Cyp51	Cytochrome P450, subfamily 51	GD 12.5−19.5	Down	−1.81	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Dab2	Disabled homolog 2 (Drosophila)	GD 12−19	Up	0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Daf1	Decay accelerating factor 1	GD 12−19	Up	0.19 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Dbi	Diazepam binding inhibitor	GD 12−19	Down	−0.38 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Dbi	Diazepam binding inhibitor	GD 12.5−19.5	Down	−1.28	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Dcc	Deleted in colorectal carcinoma	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Ddc	Dopa decarboxylase	GD 12−19	Down	−1.14 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ddc	Dopa decarboxylase	GD 18 for 18 hr	Down after 18 hr	−1.38	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Ddc	Dopa decarboxylase	GD 12.5−19.5	Down	−1.44	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Ddit4	DNA-damage-inducible transcript 4	GD 12−19	Down	−1.02 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ddit4	DNA-damage-inducible transcript 4	GD 18 for 18 hr	Down after 18 hr	−1.57	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Ddt	D-dopachrome tautomerase	GD 12.5−19.5	Down	−1.22	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Decr1	2,4-dienoyl CoA reductase 1, mitochondrial	GD 12−19	Down	−0.21 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Dhcr7	7-dehydrocholesterol reductase	GD 12−19	Down	−0.73 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Dhcr7	7-dehydrocholesterol reductase	GD 19 for 6 hr	Down after 6 hr	−1.34	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Dhcr7	7-dehydrocholesterol reductase	GD 18−19 for 18 hr	Down after 18 hr	−1.18	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Dnm3	Dynamin 3	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Dusp1	Dual specificity phosphatase 1	GD 19 for 3 hr	Up after 3 hr	0.91	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Dusp6	Dual specificity phosphatase 6	GD 12−19	Up	0.39 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Dusp6	Dual specificity phosphatase 6	GD 19 for 3 hr	Up after 3 hr	1.28	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Ebp	Phenylalkylamine Ca^{2+} antagonist (emopamil) binding protein	GD 12−19	Down	−0.64 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Echs1	Enoyl Coenzyme A hydratase, short chain 1, mitochondrial	GD 12−19	Down	−0.18 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Egr1	Early growth response 1	GD 12−19	Up	0.77 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Egr2	Early growth response 2	GD 19 for 1 hr	Up after 1 hr	1.93	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Egr2	Early growth response 2	GD 19 for 3 hr	Up after 3 hr	1.53	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Elovl5	ELOVL family member 5, elongation of long chain fatty acids (yeast)	GD 12−19	Down	−0.17 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Elovl6	ELOVL family member 6, elongation of long chain fatty acids (yeast)	GD 12−19	Down	−0.40 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Emp3	Epithelial membrane protein 3	GD 12.5−19.5	Down	−1.24	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Eno1	Enolase 1, alpha non-neuron	GD 12.5−19.5	Down	−1.63	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Enpep	Glutamyl aminopeptidase	GD 12−19	Up	0.48 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Entpd5	Ectonucleoside triphosphate diphosphohydrolase 5	GD 12−19	Down	−0.52 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Epas1	Endothelial PAS domain protein 1	GD 12−19	Down	−0.21 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ephx1	Epoxide hydrolase 1, microsomal	GD 12−19	Down	−0.57 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Erbb2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)	GD 12.5−17.5	Up	1.26	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Etfdh	Electron-transferring-flavoprotein dehydrogenase	GD 12−19	Down	−0.39 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ezr	Ezrin	GD 12−19	Up	0.20 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ezr	Ezrin	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
F10	Coagulation factor X	GD 12−19	Down	−0.51 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fabp3	Fatty acid binding protein 3	GD 12−19	Down	−0.49 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fabp3	Fatty acid binding protein 3	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Fabp3	Fatty acid binding protein 3	GD 19 for 3 hr	Down after 3 hr	−0.78	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fabp3	Fatty acid binding protein 3	GD 19 for 6 hr	Down after 6 hr	−1.68	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fabp3	Fatty acid binding protein 3	GD 18−19 for 18 hr	Down after 18 hr	−1.09	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fabp5	Fatty acid binding protein 5, epidermal	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Fabp6	Fatty acid binding protein 6, ileal (gastrotropin)	GD 12−19	Down	−0.23 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fads1	Fatty acid desaturase 1	GD 12−19	Down	−0.80 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fads1	Fatty acid desaturase 1	GD 12.5−15.5	Up	1.42	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fads1	Fatty acid desaturase 1	GD 12.5−19.5	Down	1.47	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fads2	Fatty acid desaturase 2	GD 12−19	Down	−0.42 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fat1	FAT tumor suppressor homolog 1 (Drosophila)	GD 12.5−15.5	Down	−1.32	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Fbp2	Fructose-1,6-bisphosphatase 2	GD 12−19	Up	0.28 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fdft1	Farnesyl diphosphate farnesyl transferase 1	GD 12−19	Down	−0.58 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fdft1	Farnesyl diphosphate farnesyl transferase 1	GD 12.5−19.5	Down	−1.40	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fdps	Farensyl diphosphate synthase	GD 12−19	Down	−0.73 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fdps	Farensyl diphosphate synthase	GD 12.5−17.5	Down	−1.49	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fdps	Farensyl diphosphate synthase	GD 12.5−19.5	Down	−1.41	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fdx1	Ferredoxin 1	GD 12−19	Down	−1.65 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fdx1	Ferredoxin 1	GD 18 for 18 hr	Down after 18 hr	−2.53	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fdx1	Ferredoxin 1	GD 12.5−17.5	Down	−2.06	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fdx1	Ferredoxin 1	GD 12.5−19.5	Down	−2.97	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fdxr	Ferredoxin reductase	GD 12−19	Down	−0.37 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fdxr	Ferredoxin reductase	GD 12.5−17.5	Down	−1.41	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Fgfr4	Fibroblast growth factor receptor 4	GD 12−19	Down	−0.19 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Folr1	Folate receptor 1 (adult)	GD 12−19	Down	−0.48 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fos	FBJ murine osteosarcoma viral oncogene homolog	GD 19 for 1 hr	Up after 1 hr	3.28	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fos	FBJ murine osteosarcoma viral oncogene homolog	GD 19 for 3 hr	Up after 3 hr	2.70	$p < 0.05$ (ANOVA)	Thompson et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Frag1	FGF receptor activating protein 1	GD 12−19	Down	−0.48 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Frag1	FGF receptor activating protein 1	GD 18 for 18 hr	Down after 18 hr	−0.65	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fthfd	Formyltetrahydro-folate dehydrogenase	GD 12−19	Down	−1.03 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Fthfd	Formyltetrahydro-folate dehydrogenase	GD 19 for 6 hr	Down after 6 hr	−0.98	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fthfd	Formyltetrahydro-folate dehydrogenase	GD 18−19 for 18 hr	Down after 18 hr	−0.83	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Fzd2	Frizzled homolog 2 (Drosophila)	GD 19 for 3 hr	Down after 3 hr	−0.7	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Gaa	Glucosidase, alpha, acid	GD 12−19	Down	−0.30 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ggtl3	Gamma-glutamyltransferase-like 3	GD 12−19	Down	−0.32 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Gja1	Gap junction membrane channel protein alpha 1	GD 12−19	Down	−0.36 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Glrx1	Glutaredoxin 1 (thioltransferase)	GD 12−19	Down	−0.20 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Gnrhr	Gonadotropin releasing hormone receptor	GD 19 for 3 hr	Up after 3 hr	1.38	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Gnrhr	Gonadotropin releasing hormone receptor	GD 19 for 6 hr	Up after 6 hr	2.03	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Gpsn2	Glycoprotein, synaptic 2	GD 12−19	Down	−0.42 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Grb14	Growth factor receptor bound protein 14	GD 12−19	Up	0.68 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Grb14	Growth factor receptor bound protein 14	GD 19 for 6 hr	Up after 6 hr	1.78	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Grb14	Growth factor receptor bound protein 14	GD 18−19 for 18 hr	Up after 18 hr	0.93	$p < 0.05$ (ANOVA)	Thompson et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Grina	Glutamate receptor, ionotropic, N-methyl D-aspartate-associated protein 1 (glutamate binding)	GD 12.5−15.5	Up	1.59	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Gsta2	Glutathione-S-transferase, alpha type2	GD 12.5−17.5	Down	−1.48	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Gsta2	Glutathione-S-transferase, alpha type2	GD 12.5−19.5	Down	−2.23	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Gsta3	Glutathione S-transferase A3	GD 12−19	Down	−0.96 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Gsta3	Glutathione S-transferase A3	GD 12.5−17.5	Down	−1.75	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Gsta3	Glutathione S-transferase A3	GD 12.5−19.5	Down	−2.63	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Gstm2	Glutathione S-transferase, mu 2	GD 12−19	Down	−0.42 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Gstm2	Glutathione S-transferase, mu 2	GD 12−21	Up at GD 21	$\geq$2	2-fold	Shultz et al., 2001
Gstm2	Glutathione S-transferase, mu 2	GD 18−19 for 18 hr	Down after 18 hr	−0.47	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Gsto1	Glutathione S-transferase omega 1	GD 12−19	Down	−0.42 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Gstp1	Glutathione-S-transferase, pi 1	GD 12.5−15.5	Up	1.34	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Hao2	Hydroxyacid oxidase 2 (long chain)	GD 12−19	Down	−0.58 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hmgcr	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	GD 12−19	Down	−0.47 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hmgcr	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	GD 12.5−19.5	Down	−1.83	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Hmgcs1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1	GD 12−19	Down	−1.03 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Hmgcs1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1	GD 12.5−17.5	Down	−1.72	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Hmgcs1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1	GD 12.5−19.5	Down	−1.87	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Hmox1	Heme oxygenase (decycling) 1	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hpgd	Hydroxyprostaglandin dehydrogenase 15 (NAD)	GD 12−19	Down	−0.46 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hprt	Hypoxanthine guanine phosphoribosyl transferase	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Hrasls3	HRAS like suppressor 3	GD 12−19	Down	−0.45 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hsd11b2	Hydroxysteroid (11-beta) dehydrogenase 2	GD 19 for 6 hr	Down after 6 hr	−1.16	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Hsd17b3	Hydroxysteroid (17-beta) dehydrogenase 3	GD 12−19	Up	0.28 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hsd17b7	Hydroxysteroid (17-beta) dehydrogenase 7	GD 12−19	Down	−0.32 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hsd3b1_predicted	Hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted)	GD 12−19	Down	−0.50 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted)	GD 18 for 18 hr	Down after 18 hr	−0.7	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Hspb7	Heat shock 27kD protein family, member 7 (cardiovascular)	GD 12−19	Up	0.41 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Idh1	Isocitrate dehydrogenase 1 (NADP+), soluble	GD 12−19	Down	−0.52 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Idh1	Isocitrate dehydrogenase 1 (NADP+), soluble	GD 18 for 18 hr	Down after 18 hr	−0.67	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Idi1	Isopentenyl-diphosphate delta isomerase	GD 12−19	Down	−0.85 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Idi1	Isopentenyl-diphosphate delta isomerase	GD 12.5−17.5	Down	−1.57	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Igfbp2	Insulin-like growth factor binding protein 2	GD 12−19	Down	−0.39 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Igfbp3	Insulin-like growth factor binding protein 3	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Il6st	Interleukin 6 signal transducer	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Ifitm2	Interferon induced transmembrane protein 2	GD 12.5−17.5	Down	−1.11	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Inha	Inhibin alpha	GD 12−19	Down	−1.00 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Inha	Inhibin alpha	GD 12.5−19.5	Down	−1.64	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Insig1	Insulin induced gene 1	GD 12−19	Down	−0.77 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Insl3	Insulin-like 3	GD 12−19	Down	−1.56 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
interim symbol: *Loc314323*	Interim full name: transporter	GD 12−19	Down	−0.35 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
interim symbol: *Ratsg2*	Interim name: *Ratsg2*	GD 12−19	Down	−0.13 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Kcnj8	Potassium inwardly-rectifying channel, subfamily J, member 8	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Khk	Ketohexokinase	GD 12.5−17.5	Up	1.30	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Kit	V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Krt2-8	Keratin complex 2, basic, gene 8	GD 12−19	Up	0.28 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Ldha	Lactate dehydro-genase A	GD 12.5−19.5	Down	−1.30	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Ldlr	Low density lipoprotein receptor	GD 12−19	Down	−0.79 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ldlr	Low density lipoprotein receptor	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Lhcgr	Luteinizing hormone/choriogonadotropin receptor	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Lhcgr	Luteinizing hormone/choriogonadotropin receptor	GD 19 for 6 hr	Down after 6 hr	−1.00	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Lhcgr	Luteinizing hormone/choriogonadotropin receptor	GD 18−19 for 18 hr	Down after 18 hr	−1.51	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Lhcgr	Luteinizing hormone/choriogonadotropin receptor	GD 12−19	Down	−1.39 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Limk1	LIM motif-containing protein kinase 1	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Lnk	Linker of T-cell receptor pathways	GD 19 for 3 hr	Up after 3 hr	1.17	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Lr8	LR8 protein	GD 12−19	Down	−0.45 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Lss	Lanosterol synthase	GD 12−19	Down	−0.48 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Mapk1	Mitogen activated protein kinase 1	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Marcks	Myristoylated alanine rich protein kinase C substrate	GD 12−19	Up at GD 19	≥2	2-fold	Shultz et al., 2001
Mdk	Midkine	GD 12−19	Up	0.20 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Me1	Malic enzyme 1, NADP(+)-dependent, cytosolic	GD 12−19	Down	−0.67 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Me1	Malic enzyme 1, NADP(+)-dependent, cytosolic	GD 12.5−17.5	Down	−1.36	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Men1	Multiple endocrine neoplasia 1	GD 12.5−15.5	Down	−1.17	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Mgat1	Mannoside acetylglucosaminyltransferase 1	GD 12−19	Up	0.28 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Mgp	Matrix Gla protein	GD 19 for 6 hr	Up after 6 hr	1.66	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Mgst1	Microsomal glutathione S-transferase 1	GD 12−19	Down	−0.36 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Mgst1	Microsomal glutathione S-transferase 1	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Mir16	Membrane interacting protein of RGS16	GD 12−19	Down	−0.56 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Mlxipl	MLX interacting protein-like	GD 12−19	Down	−0.31 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Mmp2	Matrix metallopeptidase 2	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Mtus1	Mitochondrial tumor suppressor 1	GD 19 for 3 hr	Up after 3 hr	0.67	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Mtus1	Mitochondrial tumor suppressor 1	GD 19 for 6 hr	Up after 6 hr	0.55	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Mvd	Mevalonate (diphospho) decarboxylase	GD 12−19	Down	−0.41 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Myd116	Myeloid differentiation primary response gene 116	GD 19 for 3 hr	Up after 3 hr	0.58	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Myh6	Myosin, heavy polypeptide 6, cardiac muscle, alpha	GD 12−19	Down	−0.72 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Myh6	Myosin, heavy polypeptide 6, cardiac muscle, alpha	GD 18−19 for 18 hr	Down after 18 hr	−1.52	$p < 0.05$ (ANOVA)	Thompson et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Myh6	Myosin, heavy polypeptide 6, cardiac muscle, alpha	GD 12.5−19.5	Down	−1.64	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Myom2	Myomesin 2	GD 12−19	Up	0.64 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Myrip	Myosin VIIA and Rab interacting protein	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nalp6	NACHT, leucine rich repeat and PYD containing 6	GD 12−19	Up	0.45 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nexn	Nexilin	GD 12−19	Up	0.26 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nf1	Neurofibromatosis 1	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Nfil3	Nuclear factor, interleukin 3 regulated	GD 12−19	Up	0.31 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nfkbia	Nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha	GD 19 for 3 hr	Up after 3 hr	0.79	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Npc2	Niemann pick type C2	GD 12−19	Down	−0.26 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nppc	Natriuretic peptide precursor type C	GD 12−19	Down	−0.56 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nr0b1	Nuclear receptor subfamily 0, group B, member 1	GD 12−19	Down	−0.37 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nr0b1	Nuclear receptor subfamily 0, group B, member 1	GD 12.5−19.5	Down	−1.15	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	GD 12−19	Up	0.3 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	GD 19 for 3 hr	Up after 3 hr	1.83	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Nr4a3	Nuclear receptor subfamily 4, group A, member 3	GD 19 for 3 hr	Up after 3 hr	2.25	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Nr5a1	Nr5a1 nuclear receptor subfamily 5, group A, member 1	GD 12.5−19.5	Down	−1.18	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Ntf3	Neurotrophin 3	GD 12.5−17.5	Up	1.34	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Okl38	Pregnancy-induced growth inhibitor	GD 12−19	Down	−0.33 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Olfm1	Olfactomedin 1	GD 12−19	Down	−0.14 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
P2ry14	Purinergic receptor P2Y, G-protein coupled, 14	GD 12−19	Down	−0.37 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Park7	Parkinson disease (autosomal recessive, early onset) 7	GD 12.5−17.5	Down	−1.32	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Pawr	PRKC, apoptosis, WT1, regulator	GD 19 for 3 hr	Up after 3 hr	1.02	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Pcna	Proliferating cell nuclear antigen	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Pcyt2	Phosphate cytidylyltransferase 2, ethanolamine	GD 12−19	Down	−0.20 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Pdap1	PDGFA associated protein 1	GD 12−21	Up at GD 21	≥2	2-fold	Shultz et al., 2001
Pdyn	Prodynorphin	GD 12−19	Down	−1.06 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Pebp1	Phosphatidylethanolamine binding protein 1	GD 12−19	Down	−0.36 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Pebp1	Phosphatidylethanolamine binding protein 1	GD 12.5−19.5	Down	−1.67	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Penk1	Proenkephalin 1	GD 12.5−17.5	Down	−1.41	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Penk1	Proenkephalin 1	GD 12.5−19.5	Down	−1.86	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Pfkp	Phosphofructokinase, platelet	GD 12.5−19.5	Down	−1.41	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Pgam1	Phosphoglycerate mutase 1	GD 12.5−19.5	Down	−1.26	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Pgk1	Phosphoglycerate kinase 1	GD 12.5−19.5	Down	−1.25	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Phb	Prohibitin	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Phb	Prohibitin	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Phyh	Phytanoyl-CoA hydroxylase	GD 19 for 6 hr	Down after 6 hr	−1.02	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Plat	Plasminogen activator, tissue	GD 12−19	Up at GD 19	≥2	2-fold	Shultz et al., 2001
Plaur	Plasminogen activator, urokinase receptor	GD 19 for 3 hr	Up after 3 hr	0.86	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Pmp22	Peripheral myelin protein 22	GD 12−19	Up at GD 19	≥2	2-fold	Shultz et al., 2001
Pmp22	Peripheral myelin protein 22	GD 19 for 3 hr	Down after 3 hr	−0.75	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Pmp22	Peripheral myelin protein 22	GD 19 for 6 hr	Down after 6 hr	−0.59	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Pnliprp2	Pancreatic lipase-related protein 2	GD 12−19	Down	−0.28 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Por	P450 (cytochrome) oxidoreductase	GD 12−19	Down	−0.64 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Por	P450 (cytochrome) oxidoreductase	GD 12.5−19.5	Down	−1.39	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Ppib	Peptidylprolyl isomerase B	GD 12.5−17.5	Down	−1.21	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Ppp1cb	Protein phosphatase 1, catalytic subunit, beta isoform	GD 12.5−17.5	Down	−1.37	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Prdx3	Peroxiredoxin 3	GD 12−19	Down	−0.53 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Prdx3	Peroxiredoxin 3	GD 18−19 for 18 hr	Down after 18 hr	−0.86	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Prdx3	Peroxiredoxin 3	GD 12.5−19.5	Down	−1.63	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Prg1	Plasticity related gene 1	GD 12−19	Down	−0.97 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Prkar2b	Protein kinase, cAMP dependent regulatory, type II beta	GD 12−19	Down	−0.33 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Prkcbp1	Protein kinase C binding protein 1	GD 12−19	Up	0.32 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Prlr	Prolactin receptor	GD 12−19	Down	−1.02 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ptma	Prothymosin alpha	GD 12−19	Down at GD 19	$-\geq 2$	2-fold	Shultz et al., 2001
Ptp4a1	Protein tyrosine phosphatase 4a1	GD 12−21	Up at GD 21	≥ 2	2-fold	Shultz et al., 2001
PVR	Poliovirus receptor	GD 19 for 3 hr	Up after 3 hr	1.26	$p < 0.05$ (ANOVA)	Thompson et al., 2005
PVR	Poliovirus receptor	GD 19 for 6 hr	Up after 6 hr	0.92	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Rabep2	Rabaptin, RAB GTPase binding effector protein 2	GD 19 for 3 hr	Down after 3 hr	−0.48	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Rasd1	RAS, dexamethasone-induced 1	GD 12−19	Down	−0.52 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Rln1	Relaxin 1	GD 12−19	Down	−0.36 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Rnh1	Ribonuclease/angiogenin inhibitor 1	GD 12.5−17.5	Down	−1.20	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rpa2	Replication protein A2	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Rpl13	Ribosomal protein L13	GD 12.5−15.5	Up	1.17	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rpl32	Ribosomal protein L32	GD 12.5−19.5	Up	1.13	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rpl37	Ribosomal protein L37	GD 12.5−19.5	Up	1.13	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rpl36a	Large subunit ribosomal protein L36a	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Rpl36a	Large subunit ribosomal protein L36a	GD 12.5−15.5	Up	1.22	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rpn2	Ribophorin II	GD 12.5−19.5	Down	−1.19	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rps13	Ribosomal protein S13	GD 12.5−15.5	Up	1.30	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rps17	Ribosomal protein S17	GD 12.5−19.5	Up	1.25	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rps19	Ribosomal protein S19	GD 12.5−17.5	Up	1.25	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Rps29	Ribosomal protein S29	GD 12.5−19.5	Down	−1.13	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Sc4mol	Sterol-C4-methyl oxidase-like	GD 12−19	Down	−1.02 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Sc4mol	Sterol-C4-methyl oxidase-like	GD 12.5−17.5	Down	−1.82	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Sc4mol	Sterol-C4-methyl oxidase-like	GD 12.5−19.5	Down	−2.36	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Sc5d	Sterol-C5-desaturase (fungal ERG3, delta-5-desaturase) homolog (*S. cerevisae*)	GD 12−19	Down	−0.32 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Scarb1	Scavenger receptor class B, member 1	GD 12−19	Down	−1.91 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Scarb1	Scavenger receptor class B, member 1	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Scarb1	Scavenger receptor class B, member 1	GD 19 for 6 hr	Down after 6 hr	−1.60	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Scarb1	Scavenger receptor class B, member 1	GD 18−19 for 18 hr	Down after 18 hr	−2.72	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Scarb1	Scavenger receptor class B, member 1	GD 12.5−17.5	Down	−2.23	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Scarb1	Scavenger receptor class B, member 1	GD 12.5−19.5	Down	−2.85	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Scd1	Stearoyl-Coenzyme A desaturase 1	GD 12−19	Down	−0.58 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Scn3b	Sodium channel, voltage-gated, type III, beta	GD 19 for 6 hr	Up after 6 hr	1.49	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Scp2	Sterol carrier protein 2	GD 12−19	Down	−0.17 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Scp2	Sterol carrier protein 2	GD 12.5−19.5	Down	−1.24	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Sdf4	Stromal cell derived factor 4	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Sepp1	Selenoprotein P, plasma, 1	GD 12−19	Down	−0.45 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Serpine1	Serine (or cysteine) peptidase inhibitor, clade E, member 1	GD 19 for 3 hr	Up after 3 hr	1.56	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Serpinh1	Serine (or cysteine) peptidase inhibitor, clade H, member 1	GD 12.5−15.5	Down	−1.32	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Sgk	Serum/glucocorticoid regulated kinase	GD 12−19	Down	−0.45 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Slc3a2	Solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	GD 12−19	Down	−0.48 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Slc12a2	Solute carrier family 12 (sodium/potassium/chloride transporters), member 2	GD 12.5−17.5	Down	−1.39	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Slc16a6	Solute carrier family 16 (monocarboxylic acid transporters), member 6	GD 12−19	Down	−0.38 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Slc25a1	Solute carrier family 25, member 1	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Slc25a20	Solute carrier family 25 (mitochondrial carnitine/acylcarnitine translocase), member 20	GD 12−19	Down	−0.23 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Slc7a8	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 8	GD 12.5−17.5	Down	−1.82	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Slc7a8	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 8	GD 12.5−19.5	Down	−2.18	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Smpx	Small muscle protein, X-linked	GD 12−19	Up	0.21 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Sod2	Superoxide dismutase 2, mitochondrial	GD 12−19	Down	−0.51 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Sod3	Superoxide dismutase 3, extracellular	GD 12−19	Down	−0.33 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Sqle	Squalene epoxidase	GD 12−19	Down	−0.59 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Sqle	Squalene epoxidase	GD 18 for 18 hr	Down after 18 hr	−1.26	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Ssr4	Signal sequence receptor 4	GD 12−19	Down	−0.23 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Ssrp1	Structure specific recognition protein 1	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Star	Steroidogenic acute regulatory protein	GD 12−19	Down	−2.45 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Star	Steroidogenic acute regulatory protein	GD 18−19 for 18 hr	Down after 18 hr	−2.33	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Star	Steroidogenic acute regulatory protein	GD 12.5−17.5	Down	−2.19	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Star	Steroidogenic acute regulatory protein	GD 12.5−19.5	Down	−2.53	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Stc1	Stanniocalcin 1	GD 12−19	Up	0.98 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Stc1	Stanniocalcin 1	GD 19 for 6 hr	Up after 6 hr	1.61	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Stc2	Stanniocalcin 2	GD 12−19	Down	−1.18 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Stc2	Stanniocalcin 2	GD 12.5−19.5	Down	−1.59	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Sts	Steroid sulfatase	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Suclg1	Succinate-CoA ligase, GDP-forming, alpha subunit	GD 12.5−19.5	Down	−1.21	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Svs5	Seminal vesicle secretion 5	GD 12−19	Down	−3.75 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Svs5	Seminal vesicle secretion 5	GD 18−19 for 18 hr	Down after 18 hr	−3.36	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Svs5	Seminal vesicle secretion 5	GD 12.5−17.5	Down	−5.89	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Svs5	Seminal vesicle secretion 5	GD 12.5−19.5	Down	−3.75	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Syngr1	Synaptogyrin 1	GD 12−19	Down	−0.16 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tcf1	Transcription factor 1	GD 12−19	Down	−0.14 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tcf21	Transcription factor 21	GD 12−19	Up	0.17 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tec	Tec protein tyrosine kinase	GD 19 for 3 hr	Up after 3 hr	0.69	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Testin	Testin gene	GD 12−19	Up	0.59 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tfrc	Transferrin receptor	GD 12−19	Down	−0.23 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tgfb3	Transforming growth factor, beta 3	GD 12−19	Down at GD 19	−≥2	2-fold	Shultz et al., 2001
Timp1	Tissue inhibitor of metallopeptidase 1	GD 19 for 6 hr	Up after 6 hr	1.04	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Timp3	Tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory)	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Tkt	Transketolase	GD 12.5−17.5	Down	−1.19	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Tkt	Transketolase	GD 12.5−19.5	Down	−1.28	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Tmed10	Transmembrane emp24-like trafficking protein 10 (yeast)	GD 12.5−19.5	Down	−1.20	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Tnfrsf12a	Tumor necrosis factor receptor superfamily, member 12a	GD 19 for 6 hr	Up after 6 hr	1.34	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Tnni1	Troponin I, skeletal, slow 1	GD 12−19	Up	0.33 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tnni3	Troponin I type 3 (cardiac)	GD 12−19	Up	0.26 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tnnt2	Troponin T2, cardiac	GD 12−19	Up	0.77 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Tpi1	Triosephosphate isomerase 1	GD 12−19	Down	−0.24 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tpm1	Tropomyosin 1, alpha	GD 12−19	Up	0.36 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tpm1	Tropomyosin 1, alpha	GD 19 for 6 hr	Up after 6 hr	1.04	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Tsc22d1	TSC22 domain family, member 1	GD 12.5−19.5	Down	−1.34	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Tsn	Translin	GD 12.5−17.5	Up	1.54	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Tst	Thiosulfate sulfurtransferase	GD 12−19	Down	−0.33 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Tuba1	Tubulin, alpha 1	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Tuba1	Tubulin, alpha 1	GD 12.5−19.5	Down	−1.26	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Txn1	Thioredoxin 1	GD 18 for 18 hr	Down after 18 hr	−0.62	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Txnl1	Thioredoxin-like 1	GD 12.5−15.5	Up	1.20	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Uba52	Ubiquitin A-52 residue ribosomal protein fusion product 1	GD 12.5−19.5	Up	1.10	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Unc5b	Unc-5 homolog B (*C. elegans*)	GD 12−21	Down at GD 21	−≥2	2-fold	Shultz et al., 2001
Vapa	VAMP (vesicle-associated membrane protein)-associated protein A	GD 12.5−19.5	Down	−1.37	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Vcam1	Vascular cell adhesion molecule 1	GD 12−19	Down	−0.63 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Vdac1	Voltage-dependent anion channel 1	GD 12.5−19.5	Down	−1.13	$p < 0.01$ (ANOVA)	Plummer et al., 2007

Table A-1 (continued)

Official gene symbol	Official gene name[f]	Exposure window	Up or down	Fold change	Cutoff used (method)	Reference
Vim	Vimentin	GD 12.5−19.5	Down	−1.60	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Vnn11	Vanin 1	GD 12−19	Down	−0.32 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Vsnl1	visinin-like 1	GD 12−19	Down	−0.62 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Ywhae	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide	GD 12.5−19.5	Down	−1.37	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Zfp36	Zinc finger protein 36	GD 19 for 1 hr	Up after 1 hr	1.79	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Zyx	Zyxin	GD 19 for 3 hr	Up after 3 hr	1.03	$p < 0.05$ (ANOVA)	Thompson et al., 2005
Not found	Listed as "Tppc" and 289920_Rn in article, and Genbank #BF400584 (Plummer, personal communication) does not match a gene name.	GD 12.5−17.5	Down	−1.39	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Not found	Listed as "Similar to mouse IAP-binding protein" and 205510_Rn in article, and Genbank #:BG378907 (Plummer, personal communication) does not match a gene name.	GD 12.5−15.5	Up	1.26	$p < 0.01$ (ANOVA)	Plummer et al., 2007
Not found	LOC499942 similar to WAP four-disulfide core domain protein 8 precursor (Putative protease inhibitor WAP8) (*Rattus norvegicus*).	GD 12−19	Down	−0.25 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005
Not found	LOC497726 hypothetical gene supported by NM_138518 (*Rattus norvegicus*). This record was discontinued.	GD 12−19	Down	−0.27 log2	$p < 0.05$ (ANOVA)	Liu et al., 2005

[a]The four studies dosed at 500 mg/kg-d DBP in the Sprague-Dawley (SD) rat.
[b]Thompson et al. (2005) and Shultz et al. (2001) dosed with DBP alone; gene expression changes for DBP were relative to vehicle control expression.

[c]Liu et al. (2005) presented microarray data for all five developmental phthalates, including DBP, since they did not find any differences in statistical significance among the five phthalates. Thus, we present the data for all five phthalates, which should be the same as for DBP.

[d]The Affy ID 1387057_at was found to be significantly down-regulated by Liu et al. (2005). This Affy ID was listed as the gene *Slc7a8* (solute carrier family 7 [cationic amino acid transporter, y+ system], member 8) at the time of their publication. As of January 2007, Affy now lists both *Slc7a8* and *Syngap1*. This probeset is apparently capable of hybridizing with two different genes. Thus, this Affy ID was not incorporated in the case study evaluation since it is not clear which gene was altered after DBP in utero exposure.

[e]The Plummer et al. (2007) data from the whole testis are included in this table. The data from microdissection of testicular regions are not presented since no other studies were comparable. Plummer et al. (2007) performed their study in the Wistar rat whereas the other three microarray studies were performed in the SD rat.

[f]Gene function and pathway information was gathered from GeneGo (www.genego.com).

ANOVA, analysis of variance; GD, gestation day; hr, hour.

Table A-2. WOE for statistically significant gene expression changes after in utero exposure to DBP from whole-rat testis reverse transcription-polymerase chain reaction (RT-PCR) studies

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Ar	Androgen receptor	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Bmp4	Bone morphogenetic protein 4	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Btg2	B-cell translocation gene 2, anti-proliferative	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−6 hr (peak ~2 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Bzrp	Benzodiazepine receptor, peripheral	500 mg/kg-d	GD 12−19	Up	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Cebpb	CCAAT/enhancer binding protein (C/EBP), beta	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA; $p < 0.05$	Liu et al., 2005
Cebpd	CCAAT/enhancer binding protein (C/EBP), delta	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−6 hr (peak ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Clu	Clusterin	500 mg/kg-d	GD 12−19	Up	ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Clu	Clusterin	500 mg/kg-d	GD 12−19	Up	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Clu	Clusterin	500 mg/kg-d	GD 12−16, 12−19, or 12−21	Up	$p < 0.05$	Shultz et al., 2001
Cxcl1	Chemokine (C-X-C motif) ligand 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−12 hr (peak at ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−19	Down	ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	50 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	100 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−16, 12−19, or 12−21	Down	$p < 0.05$	Shultz et al., 2001
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−17 and 12−18	Down at GD 18	t-test, ANOVA (one-way) with Tukey post hoc analysis; $p < 0.05$	Thompson et al., 2004
Cyp11a1	Cytochrome P450, family 11, subfamily a, polypeptide 1	500 mg/kg-d	GD 12.5−19.5	Down	One-way ANOVA followed by Bonferroni post test using GraphPad Prism; $p < 0.05$	Plummer et al., 2007
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−19	Down	Repeated measure ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−16, 12−19, or 12−21	Down at GD 19	$p < 0.05$	Shultz et al., 2001
Cyp17a1	Cytochrome P450, family 17, subfamily a, polypeptide 1	500 mg/kg-d	GD 12−17 and 12−18	Down at GD 17 and 18	t-test, ANOVA (one-way) with Tukey post hoc analysis; $p < 0.05$	Thompson et al., 2004
Daf1	Decay accelerating factor 1	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Ddc	Dopa decarboxylase	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Dusp6	Dual specificity phosphatase 6	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−12 hr (peak at ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but p value not calculated	Thompson et al., 2005
Edg3	Endothelial differentiation sphingolipid G-protein-coupled receptor 3	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−6 and 18 hr (peak ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but p value not calculated	Thompson et al., 2005
Egfr	Epidermal growth factor receptor	500 mg/kg-d	GD 12−19 and 12−21	Un-changed	t-test, $p < 0.05$	Bowman et al., 2005
Egr1	Early growth response 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−7 hr (peak ~2 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but p value not calculated	Thompson et al., 2005
Egr1	Early growth response 1	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA; $p < 0.05$	Liu et al., 2005

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Egr2	Early growth response 2	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−12 hr (peak ~2 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Fgf10	Fibroblast growth factor 10	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005
Fgfr2	Fibroblast growth factor receptor 2	500 mg/kg-d	GD 12−19 and 12−21	No stat. change	t-test, $p < 0.05$	Bowman et al., 2005
Fos	FBJ murine osteosarcoma viral oncogene homolog	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after 30 min and 6 hr (peak at 1 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Grb14	Growth factor receptor bound protein 14	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Hes6	Hairy and enhancer of split 6 (Drosophila)	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Down after 1−3 hr (peak at 3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Hsd17b3	Hydroxysteroid (17-beta) dehydrogenase 3	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Hsd17b7	Hydroxysteroid (17-beta) dehydrogenase 7	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Hsd3b1_predicted	Hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted)	500 mg/kg-d	GD 12−19	Down	ANOVA, nested design, $p < 0.05$	Barlow et al., 2003

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted) or Hsd3b1 hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	0.1 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted) or Hsd3b1 hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	1 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted) or Hsd3b1 hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	10 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted) or Hsd3b1 hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	50 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted) or Hsd3b1 hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	100 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Hsd3b1_predicted	Hsd3b1_predicted hydroxysteroid dehydrogenase-1, delta< 5 >-3-beta (predicted) or Hsd3b1 hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 1	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Ier3	Immediate early response 3	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after 1−12 hr (peak ~2 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but p value not calculated	Thompson et al., 2005
Ifrd1	Interferon-related developmental regulator 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−6 and 18 hr (peak ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but p value not calculated	Thompson et al., 2005
Igf1	Insulin-like growth factor 1	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005
Igf1	Insulin-like growth factor 1	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Igf1r	Insulin-like growth factor 1 receptor	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Igf2	Insulin-like growth factor 2	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Igfbp5	Insulin-like growth factor binding protein 5	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005
Insig1	Insulin induced gene 1	500 mg/kg-d	GD 12−19	Down	One way; and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Insl3	Insulin-like 3	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Insl3	Insulin-like 3	1000 mg/kg-d	GD 13−17 (GD 14−18 in Wilson et al., 2004 was changed to GD 13−17 to make the GD comparable to the other 7 studies)	Down	ANOVA followed by LSMEANS, $p < 0.01$ or less	Wilson et al., 2004
Insl3	Insulin-like 3	500 mg/kg-d	GD 12.5−19.5	Down	One-way ANOVA followed by Bonferroni post test using GraphPad Prism; $p < 0.05$	Plummer et al., 2007
Itgav	Integrin alpha V	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Junb	Jun-B oncogene	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	UP after ~1−12 hr (peak ~2−3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but p value not calculated	Thompson et al., 2005
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	500 mg/kg-d	GD 12−19	Down	ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	0.1 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	1 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	50 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	100 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Kit	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	500 mg/kg-d	GD 12-19	Down at GD 19	$p < 0.05$	Shultz et al., 2001
Kitl	Kit ligand	500 mg/kg-d	GD 12−19	Down	ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Ldlr	Low density lipoprotein receptor	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Lhcgr	Luteinizing hormone/choriogonadotropin receptor	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Map3k12	Mitogen activated protein kinase kinase kinase 12	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Marcks	Myristoylated alanine rich protein kinase C substrate	500 mg/kg-d	GD 12−16, 12−19, or 12−21	No stat. Change	$p < 0.05$	Shultz et al., 2001
Mgp	Matrix Gla protein	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Mmp2	Matrix metallopeptidase 2	500 mg/kg-d	GD 12−19	Up	t-test, $p < 0.05$	Bowman et al., 2005
Mmp2	Matrix metallopeptidase 2	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005
Nfil3	Nuclear factor, interleukin 3 regulated	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Nfil3	Nuclear factor, interleukin 3 regulated	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~2−24 hr (peak ~6 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Notch2	Notch gene homolog 2 (Drosophila)	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005
Npc2	Niemann Pick type C2	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Nr0b1	Nuclear receptor subfamily 0, group B, member 1	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Nr0b1	Nuclear receptor subfamily 0, group B, member 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Down at 2 hr, Up 12 hr (peak at 12 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Nr4a1	Nuclear receptor subfamily 4, group A, member 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~6 and 18 hr (peak at 12 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Nr4a3	Nuclear receptor subfamily 4, group A, member 3	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−12 hr (peak at ~3−6 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Pawr	PRKC, apoptosis, WT1, regulator	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~2−24 hr (peak ~6 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Pcna	Proliferating cell nuclear antigen	500 mg/kg-d	GD 12−16, 12−19, or 12−21	No stat. change	$p < 0.05$	Shultz et al., 2001
Prkcbp1	Protein kinase C binding protein 1	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Scarb1	Scavenger receptor class B, member 1	500 mg/kg-d	GD 12−19	Down	ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Scarb1	Scavenger receptor class B, member 1	1 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Scarb1	Scavenger receptor class B, member 1	50 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Scarb1	Scavenger receptor class B, member 1	100 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Scarb1	Scavenger receptor class B, member 1	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Scarb1	Scavenger receptor class B, member 1	500 mg/kg-d	GD 12−16, 12−19, or 12−21	Down	$p < 0.05$	Shultz et al., 2001
Scarb1	Scavenger receptor class B, member 1	500 mg/kg-d	GD 12−17 and 12−18	Down at GD 17 and 18	t-test, ANOVA (one-way) with Tukey post hoc analysis, $p < 0.05$	Thompson et al., 2004

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Sgk	Serum/glucocorticoid regulated kinase	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Down and Up; Down after 2 hr; Up after 4 and 10 hr (peak at 6 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Sostdc1	Sclerostin domain containing 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Down after 2−6 hr; Up at 18 hr (peak)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Star	Steroidogenic acute regulatory protein	500 mg/kg-d	GD 12−19	Down	Repeated measure ANOVA, nested design, $p < 0.05$	Barlow et al., 2003
Star	Steroidogenic acute regulatory protein	50 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Star	Steroidogenic acute regulatory protein	100 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Star	Steroidogenic acute regulatory protein	500 mg/kg-d	GD 12−19	Down	Dunnett's test, ANOVA (one way), $p < 0.05$	Lehmann et al., 2004
Star	Steroidogenic acute regulatory protein	500 mg/kg-d	GD 12−16, 12−19, or 12−21	Down at GD 16, 19, and 21	$p < 0.05$	Shultz et al., 2001
Star	Steroidogenic acute regulatory protein	500 mg/kg-d	GD 12−17 and 12−18	Down at GD 17 and 18	t-test, ANOVA (one-way) with Tukey post hoc analysis; $p < 0.05$	Thompson et al., 2004
Star	Steroidogenic acute regulatory protein	500 mg/kg-d	GD 12.5−19.5	Down	One-way ANOVA followed by Bonferroni post test using GraphPad Prism; $p < 0.05$	Plummer et al., 2007

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Stc1	Stanniocalcin 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~3−24 hr (peak ~6 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Svs5	Seminal vesicle secretion 5	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Tcf1	Transcription factor 1	500 mg/kg-d	GD 12−19	Down	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Tcf1	Transcription factor 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Down after 1−3 hr (peak at 1 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Testin	Testin gene	500 mg/kg-d	GD 12−19	Up	One way and two-way nested ANOVA, $p < 0.05$	Liu et al., 2005
Thbs1	Thrombospondin 1	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after 2−4 hr (peak ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Timp1	Tissue inhibitor of metalloproteinase 1	500 mg/kg-d	GD 12−21	Up	t-test, $p < 0.05$	Bowman et al., 2005
Tnfrsf12a	Tumor necrosis factor receptor superfamily, member 12a	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~1−12 hr (peak at ~3 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005
Wnt4	Wingless-related MMTV integration site 4	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after ~12 and 18 hr (peak 12 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005

Table A-2. (continued)

Official gene symbol	Official gene name*	Dose	Exposure window	Up or down	Statistical analysis method	Reference
Zfp36	Zinc finger protein 36	500 mg/kg-d	GD 19 for 30 min to 6 hr timepoints and GD 18 for 12, 18, and 24 hr time points	Up after 30 min and 6 hr and 15 and 20 hr (peak at 1 hr)	Relative expression determined using mean Ct; triplicate samples; GADPH control; SE; but *p* value not calculated	Thompson et al., 2005

*Gene function and pathway information was gathered from GeneGo (www.genego.com).

APPENDIX B

SUPPORTING TABLES AND FIGURES FOR CHAPTER 6

Appendix B contains additional tables and figures supportive of the work described in Chapter 6.

Table B-1. Differentially expressed genes that mapped to statistically significant pathways identified using the Signal to Noise Ratio (SNR) statistical filter

Gene symbol	Entrez gene ID	Gene name
Aadat	29416	Aminoadipate aminotransferase
Acadm	24158	Acetyl-Coenzyme A dehydrogenase, medium chain
Acads	64304	Acyl-Coenzyme A dehydrogenase, short chain
Acat1	25014	Acetyl-Coenzyme A acetyltransferase 1
Aco2	79250	Aconitase 2, mitochondrial
Acsl4	113976	Acyl-CoA synthetase long-chain family member 4
Akr1b4	24192	Aldo-keto reductase family 1, member B4 (aldose reductase)
Alas1	65155	Aminolevulinic acid synthase 1
Aldh1a4	29651	Aldehyde dehydrogenase family 1, subfamily A4
Aldh2	29539	Aldehyde dehydrogenase 2
Aldh6a1	81708	Aldehyde dehydrogenase family 6, subfamily A1
Aldoa	24189	Aldolase A
Aldoc	24191	Aldolase C, fructose-biphosphate
Ass	25698	Arginosuccinate synthetase
Bhmt	81508	Betaine-homocysteine methyltransferase
Chkb	29367	Choline kinase beta
Cyp11a1	29680	Cytochrome P450, family 11, subfamily a, polypeptide 1
Cyp17a1	25146	Cytochrome P450, family 17, subfamily a, polypeptide 1
Dcxr	171408	Dicarbonyl L-xylulose reductase
Ddc	24311	Dopa decarboxylase
Dhcr7	64191	7-dehydrocholesterol reductase
Ebp	117278	Phenylalkylamine Ca^{2+} antagonist (emopamil) binding protein
Ephx1	25315	Epoxide hydrolase 1
Fbp2	114508	Fructose-1,6-bisphosphatase 2
Fdft1	29580	Farnesyl diphosphate farnesyl transferase 1
Fdps	83791	Farnesyl diphosphate synthase
Fh1	24368	Fumarate hydratase 1

Table B-1. (continued)

Gene symbol	Entrez gene ID	Gene name
G6pdx	24377	Glucose-6-phosphate dehydrogenase
Gad2	24380	Glutamate decarboxylase 2
Gapdh	24383	Glyceraldehyde-3-phosphate dehydrogenase
Gatm	81660	Glycine amidinotransferase (L-arginine:glycine amidinotransferase)
Ggtl3	156275	Gamma-glutamyltransferase-like 3
Gsta2	24422	Glutathione-S-transferase, alpha type2
Gsta3	24421	Glutathione S-transferase A5
Gstm2	24424	Glutathione S-transferase, mu 2
Gstm3	81869	Glutathione S-transferase, mu type 3
Hmgcr	25675	3-hydroxy-3-methylglutaryl-Coenzyme A reductase
Hmgcs1	29637	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1
Idh1	24479	Isocitrate dehydrogenase 1 ($NADP^+$), soluble
Me1	24552	Malic enzyme 1
Mgst1	171341	Microsomal glutathione S-transferase 1
Mif	81683	Macrophage migration inhibitory factor
Mvd	81726	Mevalonate (diphospho) decarboxylase
Nos1	24598	Nitric oxide synthase 1, neuronal
Pycr2	364064	Pyrroline-5-carboxylate reductase family, member 2 (predicted)
Sqle	29230	Squalene epoxidase
Suclg1	114597	Succinate-CoA ligase, GDP-forming, alpha subunit
Tpi1	24849	Tpi1 protein

Table B-2. Genes identified using the Linear-Weighted Normalization statistical filter and mapping to the five most significant biochemical functions and /or pathways using Ingenuity

Gene symbol	Gene name
Genes mapped to integrin pathway	
F2r	Coagulation factor II (thrombin) receptor
Src	Rous sarcoma oncogene
Gng5	Guanine nucleotide binding protein (G protein), gamma 5 subunit
Gnai3	Guanine nucleotide binding protein, alpha inhibiting 3
Gng7	Guanine nucleotide binding protein, gamma 7
Mapk3	Mitogen activated protein kinase 3
Gnao1	Guanine nucleotide binding protein, alpha o
Actc1	Actin alpha cardiac 1
Camk2d	Calcium/calmodulin-dependent protein kinase II, delta
Gnaq	Guanine nucleotide binding protein
Cxcl12	Chemokine (C-X-C motif) ligand 12
Prkce	Protein kinase C, epsilon
Genes mapped to cholesterol biosynthesis/metabolism	
Hmgcs1	3-Hydroxy-3-methylglutaryl-Coenzyme A synthase 1
Hsd3b1	Hydroxyl-delta-5-steroid dehydrogenase
Dhcr7	7-Dehydrocholesterol reductase
Sqle	Squalene epoxidase
Soat1	Sterol O-acyltransferase 1
Cyp51a1	Cytochrome P450, family 51, subfamily a, polypeptide 1
Cyp27a1	Cytochrome P450, family 27, subfamily a, polypeptide 1
Hsd11b1	Hydroxysteroid 11-beta dehydrogenase 1
Hmgcr	3-Hydroxy-3-methylglutaryl-Coenzyme A reductase
Idi1	Osopentenyl-diphosphate delta isomerase
Sc4mol	Sterol-C4-methyl oxidase-like
Cyp7b1	Cytochrome P450, family 7, subfamily b, polypeptide 1

1

Table B-2. (continued)

Gene symbol	Gene name
Genes mapped to chemokine mediated signaling	
Src	Rous sarcoma oncogene
Gng5	Guanine nucleotide binding protein (G protein), gamma 5 subunit
Hmgcs1	3-Hydroxy-3-methylglutaryl-Coenzyme A synthase 1
Serpine2	Serine (or cysteine) proteinase inhibitor, clade E, member 2
Itgb5	Integrin, beta 5
Dhcr7	7-Dehydrocholesterol reductase
Gnai3	Guanine nucleotide binding protein, alpha inhibiting 3
Gng7	Guanine nucleotide binding protein, gamma 7
Sqle	Squalene epoxidase
Mapk3	Mitogen activated protein kinase 3
Gnao1	Guanine nucleotide binding protein, alpha o
Actn1	Actinin, alpha 1
Actc1	Actin alpha cardiac 1
Cav2	Caveolin 2
Cyp51a1	Cytochrome P450, family 51, subfamily a, polypeptide 1
Genes mapped to chemokine mediated signaling	
Col1a2	Procollagen, type I, alpha 2
Cfl1	Cofilin 1, non-muscle
Cav1	Caveolin 2
Hmgcr	3-Hydroxy-3-methylglutaryl-Coenzyme A reductase
Mmp2	Matrix metallopeptidase 2
Msn	Moesin
Gsk3b	Glycogen synthase kinase 3 beta
Idi1	Isopentenyl-diphosphate delta isomerase
Plat	Plasminogen activator, tissue
Sdc2	Syndecan 2
Sc4mol	Sterol-C4-methyl oxidase-like

Table B-2. (continued)

Gene symbol	Gene name
Lef1	Lymphoid enhancer binding factor 1
Vegf	Vascular endothelial growth factor
Genes mapped to glycolysis/gluconeogenesis	
Pgk1	Phosphoglycerate kinase 1
Hmgcs1	3-Hydroxy-3-methylglutaryl-Coenzyme A synthase 1
Tpi1	Triosephosphate isomerase 1
Fbp2	Fructose-1,6-bisphosphatase 2
Dhcr7	7-Dehydrocholesterol reductase
Pfkm	Phosphofructokinase, muscle
Pfkp	Phosphofructokinase, platelet
Mdh1	Malate dehydrogenase 1, NAD (soluble)
Sqle	Squalene epoxidase
Pgam1	Phosphoglycerate mutase 1
Aldoa	Aldolase A
Cyp51a1	Cytochrome P450, family 51, subfamily a, polypeptide 1
Hmgcr	3-Hydroxy-3-methylglutaryl-Coenzyme A reductase
Hk1	Hexokinase 1
Gpi	Glucose phosphate isomerase
Gapdh	Glyceraldehyde-3-phosphate dehydrogenase
Idi1	Isopentenyl-diphosphate delta isomerase
Sc4mol	Sterol-C4-methyl oxidase-like
Pfkl	Phosphofructokinase, liver

1

Table B-3. Nodes added by using Ingenuity® Pathway Analysis (IPA) software in developing the gene regulatory network for DBP

Gene symbol	Gene name
Aco1	Aconitase 1, soluble
Esrra	Estrogen-related receptor alpha
Fgf4	Fibroblast growth factor 4
Insig1	Insulin induced gene 1
Kcnj11	Potassium inwardly-rectifying channel, subfamily J, member 11
Lep	Leptin
Lnpep	Leucyl/cystinyl aminopeptidase
Nfic	Nuclear factor I/C (CCAAT-binding transcription factor)
Nme1	Non-metastatic cells 1, protein (NM23A) expressed in
Nr2f1	Nuclear receptor subfamily 2, group F, member 1
Nr5a1	Nuclear receptor subfamily 5, group A, member 1
Pld2	Phospholipase D2
Ppargc1b	Peroxisome proliferative activated receptor, gamma, coactivator 1, beta
Srebf1	Sterol regulatory element binding transcription factor 1
Srebf2	Sterol regulatory element binding transcription factor 2
Zdhhc23	Zinc finger, DHHC-type containing 23

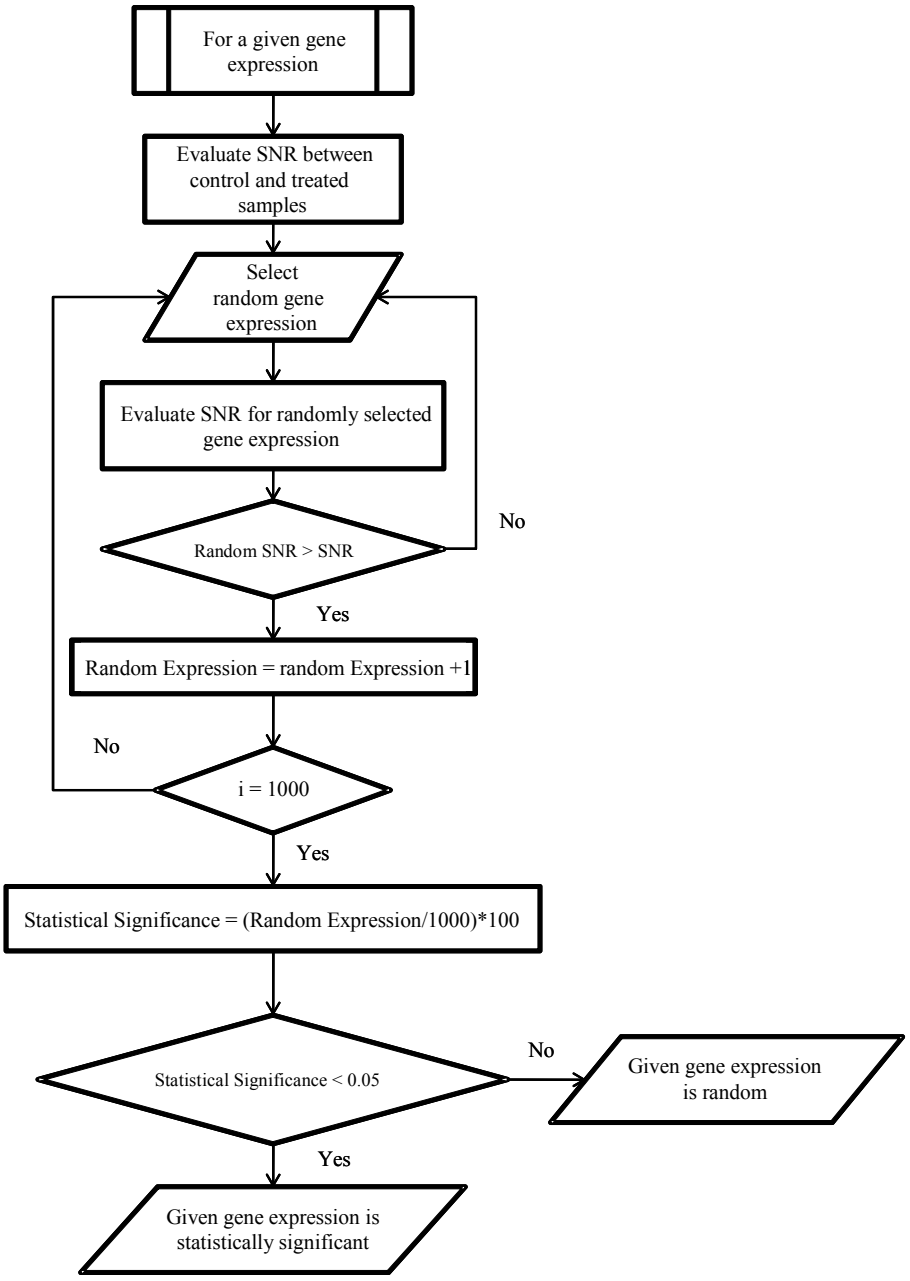

Figure B-1. Algorithm for selecting differentially expressed genes (DEGs). 1,000 random gene expressions were generated for each probe set, and then, Signal to Noise ratios (SNRs) were calculated. The ratio of the randomly generated SNR that was higher than the actual SNR determined whether individual probe set's expression was discriminating or not.

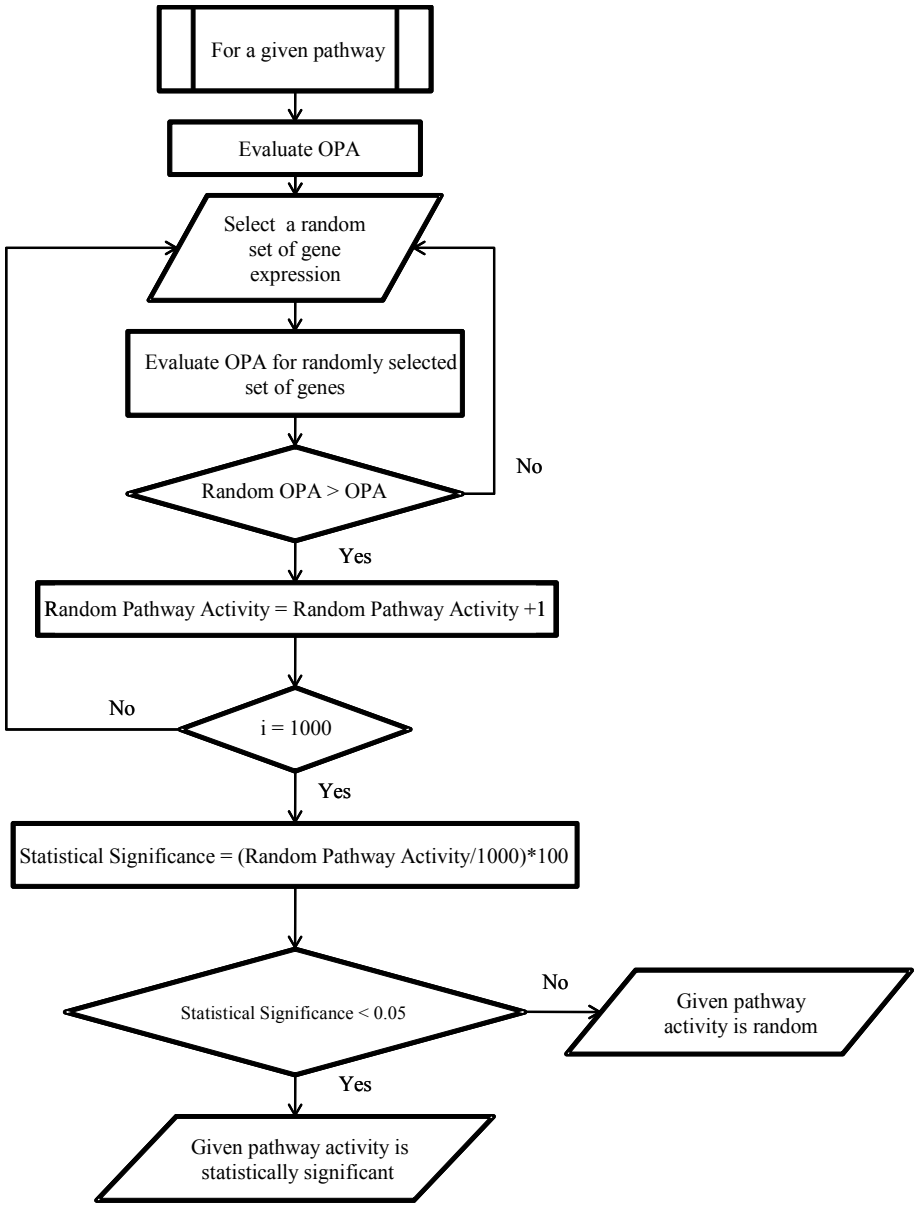

Figure B-2. Algorithm for selecting active pathways. 1,000 random sets of gene expressions were generated for each pathway, then its overall pathway activity (OPA) was calculated. The ratio of the randomly generated OPA that was higher than the actual OPA determined whether pathway activity was statistically significant.

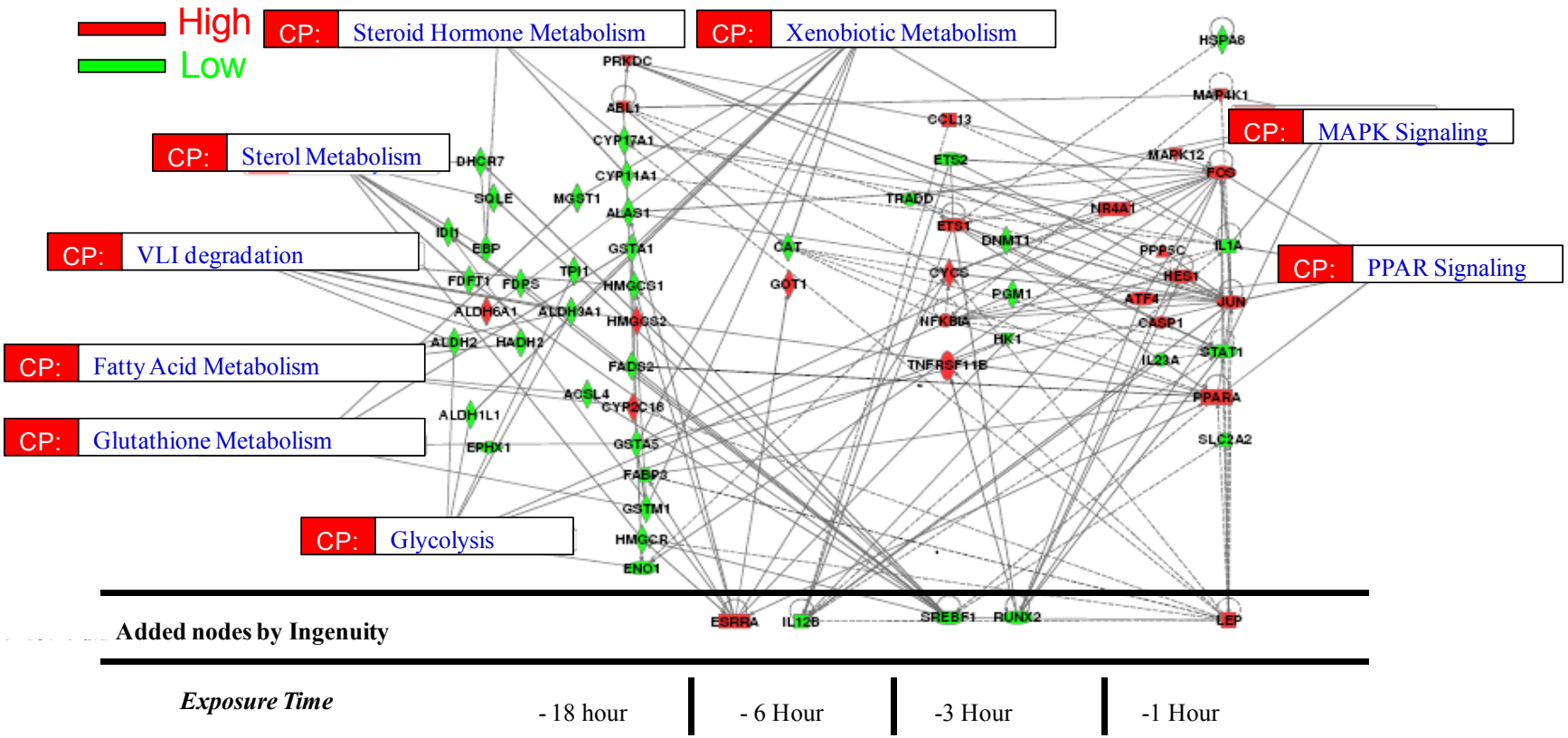

Figure B-3. Genetic regulatory network after DBP exposure created by Ingenuity® Pathway Analysis (IPA) from the informative gene list based on data from Thompson et al. (2005) The informative genes of Thompson et al. (2005) were evaluated at each time point and mapped onto a global molecular network developed from information contained in the Ingenuity® Pathways Knowledge Base.

APPENDIX C

QUALITY CONTROL AND ASSURANCE

Appendix C contains quality assurance/quality control (QA/QC) information for the work described in Chapters 5 and 6. The work described in this Appendix (C) is secondary data analysis. The studies include exploratory studies using new methods for analyzing genomic data for risk assessment purposes as well as some preliminary analyses using well-established of the raw data from two published studies.

Three projects were performed:

(1) A qualitative analysis and presentation of the 9 toxicogenomic DBP studies. No statistical analyses were performed by members of our team.

(2) In-house analysis of the raw data from Liu et al. (2005) study performed at both NHEERL, US EPA by Drs. Susan Hester and Banalata Sen, and by by collaborators, Dr. Ioannis Androulakis and Meric Ovacik, STAR Grantees at the STAR Bioinformatics Center at Rutgers/UMDNJ.

(3) New analyses of Thompson et al. (2005) data performed by collaborators, Dr. Ioannis Androulakis and Meric Ovacik, STAR Grantees at the STAR Bioinformatics Center at Rutgers/UMDNJ.

PROJECT 1

The data presented in 9 published toxicogenomic studies for DBP were compared. No additional analyses were performed. Data were entered directly into an excel spreadsheet from the published literature. Study descriptions in tables and figures were developed. The data entry process included team members entering in the data from the published articles into tables for differentially expressed genes and pathways affected. One person entered the data for a subset of genes. A second person checked the results in the table against the articles.

PROJECT 2

The data source was the DBP treatment only data from the Liu et al. (2005) study. The Liu et al. (2005) data were kindly provided by Dr. Kevin Gaido, a collaborator on this project.

The study was performed in his laboratory at The Hamner Institutes for Health Sciences (formerly CIIT). His QA statement for the collection and analysis of the data is provided below.

PROJECT 3

The data source was the Thompson et al. (2005) study. The Thompson et al. (2005) data were kindly provided by Dr. Kevin Gaido, a collaborator on this project. The study was performed in his laboratory at The Hamner Institutes for Health Sciences (formerly CIIT). His QA statement for the collection and analysis of the data is provided below.

PROJECTS 2 and 3: DATA SOURCES

The sources of the data used in the secondary analyses were the Liu et al. (2005) and Thompson et al. (2005) studies. Both of these studies were performed in the laboratory of Dr. Kevin Gaido. The QA details for the two studies are presented below. The Hamner Institute's Quality Assurance Director is Patricia O. Pomerleau, M.S., RQAP (pomerleau@thehamner.org).

A. Sample Handling Procedures

Virgin female SD outbred CD rats, 8 weeks old, were time mated. Dams were assigned to a treatment group by randomization using Provantis NT 2000 and subsequently be identified by an ear tag and cage card. Dams were kept in the Association for Assessment and Accreditation of Laboratory Animal Care International accredited animal facility at The Hamner Institute (at the time of the two studies, The Hmaner was named CIIT) in a humidity- and temperature-controlled, high-efficiency particulate-air-filtered, mass air-displacement room.

Dams were treated by gavage daily from gestation day (GD) 12−19 with corn oil (vehicle control) and dibutyl phthalate. Body weights were recorded daily before dosing (GD 12−19). The oral treatments were administered on a mg/kg-body weight basis and adjusted daily for weight changes. Animal doses were calculated through Provantis NT 2000. All calculations were checked by a second individual and recorded in the investigators' The Hamner Institute notebooks. Analytical support staff confirmed appropriate dose solutions at the beginning of the dosing period. Body weights and doses administered were recorded each day in Provantis NT 2000. Pups and dams were euthanized by carbon dioxide asphyxiation.

Fetal tissues for RIA's and RNA isolation were snap frozen in liquid nitrogen and stored at −80°C. The remaining tissues were either be embedded in optical coherence tomography and frozen or fixed in formalin for 6 to 24 hours followed by 70% ethanol and then processed and embedded in paraffin for histological examination within 48 hours. The embedded tissues were sectioned at approximately 5 microns and stained with hematoxylin and eosin. The study pathologist in consultation with the histology staff determined the gross trim, orientation, and embedding procedure for each tissue. RNA were isolated from the frozen male reproductive tract, and changes in gene expression were identified by real-time reverse transcription-polymerase chain reaction (RT-PCR) analysis (following manufacturer's protocols P/N 402876 and P/N 4304965, Applied Biosystems, Foster City, CA) and in some cases, by complementary DNA (cDNA) microarray (following manufacturers protocol PT3140, Clontech, Palo Alto, CA).

Total RNA were treated with DNase I at 37°C for 30 minutes in the presence of RNasin to remove DNA contamination before cDNA synthesis, followed by heat inactivation at 75°C for 5 minutes. Primer pairs were selected using the program Primer Express and optimized for use prior to quantification. cDNA were synthesized using random hexamers and TaqMan Reverse Transcription Reagents according to the manufacturer's suggested protocol. Real-time PCR (TaqMan) were performed on a Perkin-Elmer/Applied Biosystems 7500 Prism using TaqMan probe chemistry according to the manufacturer's instructions for quantification of relative gene expression. Relative differences among treatment groups were determined using the CT method as outlined in the Applied Biosystems protocol for reverse transcriptase(RT)-PCR. A CT value was calculated for each sample using the CT value for glyceraldehyde-3-phosphate dehydrogenase (or an appropriate housekeeping gene) to account for loading differences in the RT-PCRs.

B. ***Microarray Hybridization***

Testes from individual fetuses were homogenized in RNA Stat 60 reagent (Tel-Test, Inc., Friendswood, TX) and RNA was isolated using the RNeasy Mini Kit (Qiagen, Valencia, CA) following manufacturer's protocol. RNA integrity was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA), and optical density was measured on a NanoDrop ND 1000 (NanoDrop Technologies, Wilmington, DE). cDNA was synthesized from

2.5 or 3 µg total RNA and purified using the Affymetrix® One-Cycle Target Labeling and control reagents kit (Affymetrix, Santa Clara, CA) according to manufacturer's protocol. Equal amounts of purified cDNA per sample were used as the template for subsequent in vitro transcription reactions for complementary RNA (cRNA) amplification and biotin labeling using the Affymetrix GeneChip® IVT labeling kit (Affymetrix) included in the One-Cycle Target Labeling kit (Affymetrix). cRNA was purified and fragmented according to the protocol provided with the GeneChip® Sample Cleanup module (Affymetrix). All GeneChip® arrays were hybridized, washed, stained, and scanned using the Complete GeneChip® Instrument System according to the Affymetrix Technical Manual.

For immunocytochemistry, tissues were rapidly removed, immersed in 10% (v/v) neutral-buffered formalin for 24−48 hours, and then stored in ethanol 70% (v/v) until processed. The reproductive tissues were embedded in paraffin, sectioned at 5 µ, and processed for immunohistochemistry or stained with hematoxylin and eosin.

Experimental notes and data were entered into uniquely numbered Hamner Institute laboratory notebooks and three-ring binders along with descriptions of procedures used, according to SOP# QUA-007. Specimens (RNA and frozen tissue) were retained until analysis or discarded after a maximum of 1 year after collection. Formalin-fixed tissues, blocks, and slides were archived at the end of the study. Retention of these materials will be reassessed after 5 years.

C. *Quality Assurance*

Both QA and QC procedures are integral parts of our research program. The research was conducted under the The Hamner Institute Research Quality Standards program. These standards include (1) scientifically reviewed protocols that are administratively approved for meeting requirements in data quality, animal care, and safety regulations; (2) standardized laboratory notebooks and data recording procedures; (3) documented methods or standard operating procedures for all experimental procedures—including calibration of instruments; (4) a central managed archive for specimens and documentation; and (5) internal peer review for scientific quality of abstracts and manuscripts. The Hamner Institute QA and QC processes assessing overall study performance and records ensure that conduct of the proposed research satisfies the intended project objectives.

D. *Statistical Analysis*

RT-PCR data were analyzed using JMP statistical analysis software (SAS Institute, Cary, NC). RNA were isolated from at least 3 pups from 3 different dams for each treatment group. PCR reactions, radioimmunoassays, and protein analysis were repeated 3−5 times for each sample. Based on our experience, the number of animal replicates has the statistical power to detect a significant change in gene expression ≥20% at $p < 0.05$. The effect of treatment was analyzed using a general-linear model regression analysis. Posthoc tests were conducted when the overall analysis of variance is significant at the $p < 0.05$ level using the LS-means procedure and adjusted for multiple comparisons by Dunnett's method.

Microarray data were analyzed by a linear mixed model with SAS Microarray Solution software. Perfect-match only data were normalized to a common mean on a log2 scale, and a linear mixed model was then applied for each probe set. Restricted maximum likelihood was used for estimating the parameters for both the fixed and random effects. Significance was determined using mixed-model based F-tests ($p < 0.05$).

E. *Procedures used to Evaluate Success*

Uniquely numbered written protocols were prepared and reviewed internally prior to the start of this study. The content of a protocol includes study design, materials, laboratory methods, sample collection, handling and custody, record keeping, data analysis and statistical

procedures, animal care requirements, and safety measures. Numbered standardized laboratory notebooks and guidelines for date recording ensures completeness of data and the ability to reconstruct the study. An independent QA department manages the overall research data quality. Manuscripts describing the results of our study were prepared at the completion of each stage of this study. All manuscripts undergo a rigorous internal peer review that includes review by all authors, at least two additional PhD- level scientists, the science editor, the division manager, and the vice president for research.

PROJECT 2: DATA REVIEW, VERIFICATION, AND VALIDATION

Banalata Sen received the Liu et al. (2005) raw data files from Dr. Kevin Gaido. Two team members, Dr. Banalata Sen (National Center for Environmental Assessment, Research Triangle Park [NCEA-RTP]) and Dr. Susan Hester (National Health and Environmental Effects Research Laboratory [NHEERL]) performed the data analysis at NHEERL, RTP. Barbara Collins (collins.barbara@epa.gov) at NHEERL-RTP has agreed to serve as the Quality Assurance Manager (QAM) for the project. Dr. Hester and Sen performed analyses of the "DBP only" data that is a subset of the data presented in Liu et al. (2005). The analyses at NHEERL included statistical filtering to identify of differentially expressed genes and pathway analysis.

A. VERIFICATION OF DATA UPON RECEIPT

Upon receiving data from Kevin Gaido at the Hamner Institute, EPA NHEERL scientisits conducted a QA review of the data by gross inspection of the cel files to confirm that the data had been transmitted successfully. The scientists at the STAR Bioinformatics Center/Rutgers received the data files from Susan Euling at EPA NCEA who had received the data from Kevin Gaido at the Hamner Institute. Kevin Gaido gave permission to Susan Euling to provide the data for these analyses. A review of the data was performed by inspection of the txt files and the published data to confirm that the data had been transmitted successfully.

B. VERIFICATION OF DATA ANALYSIS CALCULATIONS

EPA NHEERL used a principal component analysis (PCA) to evaluate the within-group and across-group variance of the six samples. PCA elucidates the separation of different treatment groups and provides information about whether the data contain significant information. This was conducted using the raw data cel files in Rosetta Resolver Software. The analyses were in silico without functional validation (RT-PCR of individual genes).

The Star Bioinformatics Center also performed a principal component analysis (PCA) and displayed a 3-D plot to evaluate the within-group and across-group variance of the samples. This was conducted using the txt files in MATLAB® Software. This was an in silico analysis. The data were normalized to a zero mean and a unity standard deviation over samples. They assessed the degree of separation for Liu et al. (2005) data. A regular regular t-test and ANOVA analyses of the data were performed. The filtered data were visualized in a heatmap to determine the statistically significant subset of genes to provide a differentially expressed gene (DEG) list.

Drs. Susan Hester and Banalata Sen also performed some comparative analyses between the two outpus (above). The two independent analyses of the same dataset were contrasted with one another. Correlation plots comparing the Log10 average intensities of control samples vs. DBP treated samples was performed in order to determine the noise in both groups. Average background signal and scaling factors will be applied based on the vendor recommendations. QC plots will be made to determine the relationship between light intensity and each genechip.

PROJECT 3: DATA REVIEW, VERIFICATION, AND VALIDATION

This project analyzed the time-course data from Thompson et al. (2005) dataset to then build a regulatory network model. The STAR Center's internal QA/QC procedures are implemented and monitored by a QA official, Clifford Weisel (weisel@eohsi.rutgers.edu), at Rutgers University that reports to the National Center for Environmental Research (NCER), the granting organization for the STAR program.

A. VERIFICATION OF DATA UPON RECEIPT

Data were received from Susan Euling at EPA who had received the data from Kevin Gaido at the Hamner Institute. Kevin Gaido gave permission to Susan Euling to provide the data for these analyses. A review of the data was performed by inspection of the txt files and the published data to confirm that the data had been transmitted successfully.

B. VERIFICATION OF DATA ANALYSIS CALCULATIONS

A principal component analysis (PCA) was performed and a 3-D plot was displayed to evaluate the within-group and across-group variance of the samples. This was conducted using the txt files in MATLAB® Software. This was an in silico analysis. The data were normalized to a zero mean and a unity standard deviation over samples. They assessed the degree of separation for the Thompson et al. (2005) data. A regular regular t-test and ANOVA analyses of the data were performed. The filtered data will be visualized in a heatmap to determine the statistically significant subset of genes to provide a differentially expressed gene (DEG) list.