National Toxicology Program
Board of Scientific Counselors

November 30 – December 1, 2010

National Institute of Environmental Health Sciences
Research Triangle Park, NC

Table of Contents

I. Frequently Used Abbreviations and Acronyms

ACToR	Aggregated Computational Toxicology Resource
ADME	absorption, distribution, metabolism, and excretion
AFB1	aflatoxin B1
APWG	Assays and Pathways Working Group
AP	aromatic phosphate
AR	androgen receptor
ATP	adenosine triphosphate
ATSDR	Agency for Toxic Substances and Disease Registry
BDE	brominated diphenyl ether
BPA	bisphenol A
BPDP	tert-butylphenyl diphenyl phosphate
BSB	Biomolecular Screening Branch
BSC	Board of Scientific Counselors
CCL2	chemokine (C-C motif) ligand 2
CDER	Center for Drug Evaluation and Research
CEBS	Chemical Effects in Biological Systems
CERHR	Center for the Evaluation of Risks to Human Reproduction
CFSAN	Center for Food Safety and Applied Nutrition
CLND	chemiluminescent nitrogen detection
CP	chlorpyrifos
CPSC	Consumer Product Safety Commission
CSWG	Chemical Selection Working Group
CTD	Comparative Toxicogenomics Database
DCA	dichloroacetic acid
DERT	Division of Extramural Research and Training
DIR	Division of Intramural Research
DMSO	dimethy sulfoxide
ELSD	evaporative light scattering detection
ER	estrogen receptor
EPA	Environmental Protection Agency
ESI	electrospray ionization
FAI	free androgen index
FDA	Food and Drug Administration
FF	fresh frozen
FFPE	formalin-fixed, paraffin-embedded
FN	false negative
FP	false positive
FOIA	Freedom of Information Act
FSH	follicle stimulating hormone
GD	gestational day
GOS	Gulf oil spill
GWAS	genome-wide association studies
HHS	Health and Human Services
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
HPV	high production volume
HTS	high throughput screening
IPP	isopropylated phenol phosphate
IWG	Informatics Working Group

LBD	ligand-binding domain
LC/MS	liquid chromatography with mass spectrometry
LH	luteinizing hormone
µm	micrometer
ML	Molecular Libraries
MOU	Memorandum of Understanding
MTD	maximum tolerated dose
NAS	National Academy of Sciences
NBBS	N-butylbenzenesulfonamide
NCCT	National Center for Computational Toxicology
NCI	National Cancer Institute
NCGC	NIH Chemical Genomics Center
NCTR	Nation Center for Toxicological Research
NHANES	National Health and Nutrition Examination Survey
NICHD	National Institute of Child Health and Development
NIEHS	National Institute of Environmental Health Sciences
NIH	National Institutes of Health
NIOSH	National Institute of Occupational Safety and Health
NOAEL	no observed adverse effect level
NR	nuclear receptor
NRC	National Research Council
NTP	National Toxicology Program
OECD	Organisation for European Economic Cooperation
ORD	Office of Research and Development
OS	oxidative stress
OSHA	Occupational Safety and Health Administration
PETA	People for the Ethical Treatment of Animals
PFAA	perfluoroalkyl acids
PFOA	perfluorooctanoic acid
PFOS	perfluorooctanesulfonic acid
ppm	parts per million
QC	quality control
qHTS	quantitative HTS
qNPA	quantitative nuclease protection assay
qPCR	quantitative polymerase chain reaction
QSAR	quantitative structure-activity relationship
RoC	Report on Carcinogens
ROS	reactive oxygen species
RTK	reverse toxicokinetics
SBIR	Small Business Innovation Research
SCVCE	sperm count and vaginal cytology evaluation
SNP	single nucleotide polymorphism
STTR	Small Business Technology Transfer
TK	toxicokinetic
ToxPi	Toxicological Prioritization Index
TPP	triphenyl phosphate
TTWG	Targeted Testing Working Group
UNC	University of North Carolina at Chapel Hill
UV-DAD	ultraviolet diode array detection

II. Attendees

Members in Attendance:
Tracie Bunton, Eicarte LLC (December 1 only)
Edward Carney, Dow Chemical Company
Russell Cattley, Amgen
Elaine Faustman, University of Washington (December 1 only)
William Janzen, University of North Carolina at Chapel Hill (UNC)
Raymond Novak, Shriners Hospital for Children International (Chair)
Ruthann Rudel, Silent Spring Institute
James Sherley, Boston Biomedical Research Institute
Gina Solomon, Natural Resources Defense Council (November 30 only)
Justin Teeguarden, Pacific Northwest National Laboratory

Members not in attendance:
David Eastmond, University of California
Janan Eppig, The Jackson Laboratory
Stephen Looney, Medical College of Georgia
Mitzi Nagarkatti, University of South Carolina School of Medicine

Pending Board Members:
Miguel Fernandez, University of Texas Health Science Center at San Antonio
Nicholas Jewell, University of California Berkeley
Dana Loomis, University of Nebraska Medical Center
Richard Miller, GlaxoSmithKline
Lisa Minor, *In Vitro* Strategies, LLC
Judith Zelikoff, New York University School of Medicine (via telephone)

***Ad Hoc* Member:**
Tim Wiltshire, UNC

Other Federal Agency Staff:
Christopher Austin, NIH Chemical Genomics Center (NCGC)
Edward Bearden, Food and Drug Administration (FDA)
R. Daniel Benz, FDA
David Dix, Environmental Protection Agency (EPA)
Cherie Estill, National Institute for Occupational Safety and Health (NIOSH)
Ruili Huang, NCGC
Michael-Rock Goldsmith, EPA
Richard Judson, EPA
William Mundy, EPA
Paul Howard, FDA
David Reif, EPA
Steven Schrader, Centers for Disease Control and Prevention (CDC)
Matias Attene Ramos, NCGC
Ivan Rusyn, UNC
Imran Shah, EPA
Mark Toraason, NIOSH
John Wambaugh, EPA
Hang Wang, NCGC
Elizabeth Whelan, American Council on Science and Health

Menghang Xia, NCGC

National Institute of Environmental Health Sciences (NIEHS) Staff:

Danica Andrews	Jonathan Freedman	Scott Masten	Christina Teng
Scott Auerbach	John French	Elizabeth Maull	Kristina Thayer
Mamta Behl	Laura Fuhrman	Barry McIntyre	Raymond Tice
Linda Birnbaum	Dori Germolec	B. Alex Merrick	Molly Vallant
Jack Bishop	Xiaohong Gu	Fred Parham	Michael Waalkes
Windy Boyd	Robbin Guy	Cynthia Rider	Suramya Waidyanatha
John Bucher	Gloria Jahnke	Ruchir Shah	Nigel Walker
Xiaoqing Chang	Paul Jung	Michael Shelby	Vickie Walker
Rajendra Chhabra	Grace Kissling	Keith Shockley	Lori White
Bradley Collins	JoAnn Lewis	Robert Sills	Kristine Witt
Michael Cunningham	Ruth Lunn	Cynthia Smith	Mary Wolfe
Michael DeVito	Robin Mackar	Diane Spencer	
Paul Foster	David Malarkey	William Suk	

Public:

Nour Abdo, UNC
Neepa Choksi, Integrated Laboratory Systems, Inc (ILS)
Patrick Crockett, SRA International
Wendy Haines, ILS
Marcus Jackson, ILS
Wendelyn Jones, CropLife America
Joseph Manuppello, People for the Ethical Treatment of Animals (PETA)
Glenn Myatt, Leadscope, Inc.
Yen Low, UNC
Hirohisa Nagahori, The Hamner Institutes/Sumitomo Chemical
Maria Smith, SRA
Valerie Soldatow, UNC
Samantha Suiter, PETA
Abraham Tobia, Nufarm Americas, Inc.
Richard Woychik, The Jackson Laboratory
Fred Wright, UNC
Leah Zorrilla, ILS

November 30, 2010

III. Introductions and Welcome

The National Toxicology Program (NTP) Board of Scientific Counselors (BSC) met November 30 – December 1, 2010, in Rodbell Auditorium, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, North Carolina. Dr. Raymond Novak served as chair. He welcomed everyone to the meeting and asked BSC members and other attendees to introduce themselves. Dr. Lori White read the conflict of interest policy statement. She noted that Dr. Wiltshire, who was attending as an *ad hoc* reviewer, collaborates with Dr. Rusyn at the University of North Carolina, and as such would not comment on Dr. Rusyn's presentation. She also noted that BSC pending member Dr. Judith Zelikoff would be participating via telephone.

IV. Report of the NIEHS/NTP Director

A. Presentation

Dr. Linda Birnbaum, Director of NIEHS and NTP, welcomed attendees to the meeting and noted that it had been six months since the last BSC meeting.

In staff developments, she said the NIEHS Bethesda office is now fully staffed. Recent additions include Dr. John Balbus, Senior Adviser for Public Health, Dr. Aubrey Miller, Senior Medical Officer, and toxicologist Dr. Christopher Weis, who serves as liaison to the toxicological community. They join Legislative Liaison Mary Gant in the Bethesda office. There have also been additions to the Office of the Director team, as Bruce Androphy, J.D. was named Director of the Office of Ethics and Deputy Ethics Coordinator, Dr. Paul Jung was named Chief of Staff, and Dr. Ericka Reid was named Director of Outreach and Education.

Dr. Birnbaum announced that Dr. Rick Woychik, former CEO of the Jackson Laboratory, Bar Harbor, Maine, had been named NIEHS Deputy Director. Dr. Woychik will be leading the development of a new 2012-2016 strategic plan for NIEHS. Dr. Birnbaum also announced that Dr. Gwen Collman had been named Director of the NIEHS Division of Extramural Training and Research (DERT).

In other developments regarding permanent staff, the search for a new NIEHS Scientific Director has been re-opened, and the searches for a Clinical Director and an Executive Officer continue. Dr. Birnbaum noted that the NIEHS/NTP reorganization plan, which would establish the NTP as a separate division within NIEHS, has been signed by NIH Director Dr. Francis Collins and is being submitted to the Department of Health and Human Services for final approval.

Regarding NIEHS/NTP appropriations, she said the situation is still somewhat unknown as a result of the recent election. The President's request represents a 2.6% increase. She noted that NTP is not and has never been a line item in the NIH budget. The President's request for NIEHS Superfund research and worker training programs rose 3.1%, to over $81 million. The three House Appropriations Subcommittees that have jurisdiction over NIEHS funding have marked up their bills, but they have not been reported out. The Senate Appropriations Committee, however, has issued its report language, urging NIEHS to enhance research in endocrine disrupting chemicals and women's health, exposures related to autoimmune diseases, exposures from cosmetics and personal care products, exposures related to increased time to pregnancy, and implementation of the 2007 National Academy of Sciences (NAS) report, *Toxicity Testing in the 21st Century*. The Senate also indicated continuing interest in the Genes and Environment Initiative, as well as the Sister Study.

Dr. Birnbaum delineated the contributions to the NTP budget from its participating agencies since 2008. In 2010, the NIEHS contribution was $101.3 million, the NCTR contributed $21.6 million, and NIOSH contributed $18.3 million. In addition, NIEHS provided $1.4 million to NIOSH and $14.7 million to NCTR for NTP-related research activities at those agencies. NIEHS also funded a study at the EPA ($0.4 million) and provided $4 million to the NIH Chemical Genomics Center (NCGC) for work related to the Tox21 initiative.

She updated the BSC on continuing NIEHS/NTP activities related to the Gulf oil spill (GOS): (1) the NTP is conducting toxicological studies to identify important biological and tissue targets for crude oil fractions, (2) the Division of Intramural Research (DIR) is spearheading a longitudinal, prospective cohort study (GuLF STUDY) to assess short-and long-term health effects of exposure to oil spill, and (3) with several other NIH Institutes, DERT will fund applications to examine the impacts of the GOS on the health and quality of life of the general population residing in the Gulf Coast Region.

She noted the meetings that have taken place related to the GOS: two Institute of Medicine workshops (June 22-23, September 22-23), an interagency toxicology workshop (October 13), and an interagency workshop on federal data related to the GOS for human health (November 17).

B. Recognition of Retiring Members

Dr. Birnbaum thanked and presented certificates of appreciation to retiring BSC members Dr. Edward Carney, Dr. Russell Cattley, Mr. William Janzen, and chair Dr. Novak. Dr. Tracie Bunton was not present at this day's session, and was recognized during the December 1 proceedings.

C. BSC Discussion

Dr. James Sherley asked how connected the Congressional language was with input received from NIEHS. Dr. Birnbaum said that was a difficult question, in that NIEHS cannot go directly to Congress, but must be invited. She noted that since she had become Director she had done "quite a bit of Hill work." With her frequent testimony, she felt that the voice of NIEHS had been heard. She noted that many NIEHS stakeholders are closely aligned with Congressional colleagues or friends, and speak to them frequently. There is, however, no direct line from NIEHS to the Congress.

Dr. Novak welcomed Dr. Zelikoff, who joined the meeting by telephone.

V. Contract Concept: NTP Sperm Count and Vaginal Cytology Evaluation (SCVCE) (ACTION)

A. Presentation

Ms. JoAnn Lewis, Office of Acquisitions at the NIEHS, briefly outlined the guidelines for the BSC regarding the discussion of research concepts. She asked the BSC to review the concept for its overall value and for its scientific relevance to fulfill the program's goal of protecting public health. The specific areas to consider are scientific, technical, and programmatic significance; availability of the technology and other resources necessary to achieve the required goals; extent to which there are identified, practical, scientific or clinical uses for the anticipated results; and where pertinent, adequacy of the methodology to be used to perform the activity. The discussion should be limited to a review of the general purpose, scope, goal, and optional approaches to pursue the overall objectives. Ms. Lewis said the meeting would be closed to the public should discussions turn to the development or selection of the details of the project such

as specific technical approaches, protocol, statement of work, data format, or product specifications. If necessary, the meeting would be closed to protect free exchange of the advisory group members' opinions and avoid premature release of the details of the proposed contract or request for proposal.

Dr. Jack Bishop, NIEHS/NTP, presented the concept for a contract, which is a recompetition of an existing NTP contract for the continued evaluations of reproductive tissues obtained from rats and mice used in the NTP's 90-day toxicity studies. Specifically, the evaluations are for chemically induced changes in the number and motility of caudal sperm and in the number of testicular spermatids from male rats and mice, and in vaginal cytology used to determine the estrous cyclicity of female rats and mice.

Dr. Bishop emphasized that the tissues are being obtained from animals already being used in NTP 90-day toxicity studies, and thus additional animals are not needed for the studies. After the tissues have been evaluated and the data captured in the new data capture and analysis system, a report is prepared by the sperm count and vaginal cytology evaluation (SCVCE) contractors for presentation to the NTP. He noted that reproductive toxicity continues to be of major concern to both the public, to regulatory agencies such as EPA and FDA, and to the NTP. With these concerns, the need to continue testing chemicals for possible reproductive toxicity in both males and females is greater than ever.

He reported that the NTP has conducted SCVCEs for more than 25 years on more than 250 environmental agents. These evaluations continue to be important to the NTP for identifying agents with a potential for reproductive toxicity, for identifying which sex or sexes may be affected, and for ranking agents for further testing. He noted that the work to be performed under the new contract is the same as that currently conducted under an existing SCVCE contract.

B. BSC Discussion

Dr. Carney, lead discussant, said the proposal was straightforward, and was continuing an important activity that had been conducted for a long time. He expressed support for the concept.

Dr. Wiltshire asked whether the group was planning to examine sperm structure. Dr. Bishop said at one time the program looked at sperm morphology, but had determined that those evaluations were not particularly informative in terms of reproductive effects that might be seen in a full reproductive study of fertility or fecundity. He said the sperm samples are available should anyone wish to conduct such an analysis.

Dr. Novak asked for a motion to approve the concept. Dr. Carney moved to approve the concept and Mr. Janzen seconded the motion. The seven BSC members present approved the motion unanimously.

Review of the Biomolecular Screening Branch (BSB)

VI. Overview of the BSB and the Tox21 Initiative

A. Presentation

Dr. Raymond Tice, NIEHS/NTP BSB Chief, briefly reviewed the agenda for the BSB/Tox21 review, stressing that it would be important for the BSC to look at the totality of the effort, as opposed to concentrating on each separate entity, because the integration of the elements is the strength of the program. He pointed out that Tox21 involves the efforts of four different governmental organizations, working in harmony toward a common goal.

Dr. Tice presented the pending organizational structure of the NTP; it includes the BSB, which was established in late 2007 and became fully functional in late 2008. He related the NTP Vision for the 21st Century as drafted in 2004, which calls for toxicology to evolve from a predominantly observational science to a predominantly predictive science. To implement that vision, the NTP developed a Roadmap for the Future, which included a major new initiative to establish a high throughput screening (HTS) program with three main goals: (1) to identify mechanisms of action for further investigation, (2) to prioritize substances for further in-depth toxicological evaluation, and (3) to develop predictive models for *in vivo* biological response.

In late December 2005, NTP conducted an HTS Assays Workshop to gather information about HTS and its applications in toxicology screening. The workshop substantially guided future efforts and developments. The workshop participants, many of whom have been involved in Tox21, came from a wide range of backgrounds.

Dr. Tice provided background about the NCGC, one of the Tox21 partners. The NCGC was established in 2004 to use HTS methods to identify small molecules that can be optimized as chemical probes to study functions of genes, cells, and biochemical pathways. In 2005, NTP established collaboration with NCGC and provided an initial set of assays and a 1408-compound library for proof-of-principle studies. Screening of that library began in 2006, when NCGC also established a similar collaboration with the EPA National Center for Computational Toxicology (NCCT).

In 2007, the NAS published a report entitled *Toxicity Testing in the 21st Century: A Vision and Strategy*, which envisioned a near future in which all routine toxicity testing would be conducted *in vitro* in human cells or cell lines using HTS methodologies. Efforts would concentrate on evaluating perturbations of cellular responses in a suite of toxicity pathway assays. The report illustrated the activation of a toxicity pathway through perturbations, with a low dose allowing normal biologic function and higher doses leading to adverse health outcomes.

Dr. Tice showed a diagram illustrating the components of the NAS Vision, the elements of which have been central to the development of the Tox21 initiative. The components include: chemical characterization, toxicity testing encompassing toxicity pathways and targeted testing, and dose-response and extrapolation modeling, within a framework outlined by risk contexts and population and exposure data. He also described a diagram in the NAS report that illustrates the process of risk characterization, encompassing hazard identification and exposure assessment, including dose-response assessment, all ultimately leading to regulatory guidance.

He described the NAS report requirements for an implementation strategy, which include: (1) a comprehensive suite of *in vitro* tests, preferably based on human cells, cell lines, or components; (2) targeted animal tests to complement *in vitro* tests; (3) computational models of toxicity pathways to support application of *in vitro* test results in risk assessments; (4) infrastructure changes to support basic and applied research needed to develop the tests and pathway models; (5) validation of tests and test strategies; and (6) evidence justifying that the toxicity pathway approach is adequately predictive of adverse health outcomes to use in decision-making.

The NAS report focused on knowledge development, including the identification of toxicity pathways and multiple pathways, the nature of adversity, the impact of life stages, the effects of exposure duration, low-dose response, and human variability. The report also focused on method development, calling for the development of methods to predict metabolism, tools for chemical characterization, assays to uncover cell circuitry, assays for large-scale application, suites of assays, a strategy for human surveillance, new mathematical models for data interpretation and extrapolation, and inclusion of the concept of test strategy uncertainty.

In response to the NAS report and reflecting the fact that the NCGC, the EPA, and the NTP had already begun a collaboration in these areas, in February, 2008, a Memorandum of Understanding (MOU) entitled *High-Throughput Screening: Toxicity Pathway Profiling and Biological Interpretation of Findings* was signed by the NIEHS/NTP, the National Human Genome Research Institute (NHGRI)/NCGC, and the U.S. EPA/Office of Research and Development (ORD). The Tox21 MOU built on existing expertise, overcoming the resource limitations of a single agency. The partners agreed to collaborate on new toxicity pathway test methods, with the data ultimately to be provided to risk assessors. The key sections of the MOU delineated toxicity pathways, chemical selection, analysis and bioinformatics, outreach, scientific review (including BSC review) and governance.

Dr. Tice outlined the goals of Tox21, which are focused on achievable objectives: (1) research, develop, validate and translate innovative compound testing methods that characterize toxicity pathways; (2) identify compounds, assays, informatics tools, and targeted testing needed for the innovative testing methods; (3) prioritize compounds for more extensive toxicological evaluation; (4) identify mechanisms of compound-induced biological activity; and (5) develop predictive models for biological response in humans.

The goals of the BSB include, but are not limited to, developing (1) automated screening assays with *C. elegans*, (2) research and testing activities in medium and high throughput screening assays for rapid detection of biological activities of significance to toxicology and carcinogenesis, (3) computational tools and approaches to allow an integrated assessment of HTS endpoints and associations with findings from traditional toxicology and cancer models, and (4) assays and approaches to understand the genetic and epigenetic bases for differences in susceptibility. These activities were described in detail later in the meeting.

Dr. Tice presented a ten-year, time/risk matrix developed among the Tox21 partners shortly after the MOU was signed. The matrix has guided planning among the partners, based on level

of difficulty versus time, extending to 2018. He said in his summary during the closing session, that he would go into more detail on which goals on the matrix have been accomplished, which are still be addressed, and which are still targets for the future.

He reported that on July 19, 2010, a revised MOU was released, reflecting the expansion of the Tox21 community with the addition of the FDA. Illustrated by a chart depicting the various areas of expertise covered by the Tox21 partners, he mentioned that the FDA has added the ability to access needed human toxicological data by being a liaison to the pharmaceutical industry.

Dr. Tice previewed the rest of the talks in the Tox21 review portion of the meeting, with presentations by the agencies' points of contact, BSB representatives of the four Tox21 working groups, and reports from staff on five of the major areas of Tox21 activity specific to the NTP. He concluded by stating that the Tox21 effort is now at the cusp between testing small numbers of compounds and larger libraries, and that feedback from the BSC would be appreciated.

B. BSC Questions

Dr. Gina Solomon commented that the NTP Roadmap had been extremely influential in the thinking that went into the 2007 NAS report, contrary to Dr. Tice's assertion that the NAS had been mainly unaware of the activities that had been undertaken by NTP, EPA, and NCGC up to that point.

Dr. Mark Toraason asked why there is no division or specific program in Tox21 to assess predictivity in terms of human health, and what plans were in place to do so. Dr. Tice said subsequent presentations would provide a much clearer picture of how that issue is being addressed within Tox21. He acknowledged the difficultly assessing human health effects without human data, but that as the program has progressed, more potential partners are expressing interest, with potential contributions to the overall knowledge base. Dr. Birnbaum added that at a recent NIH Director's retreat there had been much interest in translational medicine. She said as new knowledge emerges about specific pathways and signatures, particularly in cancer research, there would be opportunities for further collaborations. Dr. Tice reminded the BSC that the NTP Host Susceptibility Branch had been incorporated into the BSB, bringing its expertise in seeking homologous pathways and genes between mouse models and human disease.

Dr. Bucher said one of the things that the NTP would like to get from the BSC's review is counsel regarding how to pull together the various Tox21 elements to achieve the goal of positively influencing human health decisions.

Dr. Sherley asked Dr. Tice to describe the actionable set of prioritized compounds involved. Dr. Tice replied that EPA had developed a prioritization strategy that had been incorporated into Tox21 efforts. He clarified references to the 10,000 (10K) compound library, stating that there is a difference between pharmaceutical and toxicological screening in efforts to identify compounds that have activity. Pharmaceutical companies ("pharma"), he said, are not as concerned as toxicologists about the extent of false negatives (FNs) or false positives (FPs) in an assay, since their focus is on identifying strong actives. In toxicology, there is more concern

about the whole breadth of activity. He said the decision had been made to test every compound in the 10K library three times in the same run, over a 14-point concentration range, with the compound located in a different plate location for each set of concentrations. This approach should help to reduce the number of inconclusive results. Identification of FPs and FNs for any one assay would depend largely on examining screening results from batteries of assays, some of which would involve related endpoints. To prioritize compounds for more comprehensive testing in lower throughput but more informative *in vitro* or *in vivo* assays, potency for a specific target, such as the estrogen receptor (ER), could be used. Another way to prioritize compounds would be based on the number of different pathways affected. Thus, compounds are triaged at various stages of the process.

VII. Tox21 Partners: U.S. Environmental Protection Agency

A. Presentation

Dr. David Dix, EPA, outlined EPA Administrator Jackson's principles for chemical reform, concentrating on the first, which states "EPA must review all chemicals against risk-based safety standards." That principle presents the daunting challenge to EPA to review *all* chemicals, of which there are thousands that may have human health effects. Dr. Dix said in the past, the tools have not been available to do so, but with the HTS offered by Tox21 and the EPA's ToxCast™ project, that is no longer the case. Administrator Jackson also mandated the EPA to "encourage innovation in green chemistry and sustainable processes," to which Tox21 and ToxCast™ contribute.

The EPA ToxCast™ Project, initiated in 2005, was designed to use bioactivity profiling, *in vitro* testing, and HTS to assess thousands of chemicals at a much lower cost and in a much shorter period of time than was possible with traditional, panel-based toxicity testing. *In silico* analysis, i.e., computational toxicology, is the second step in the process, required to make sense of the *in vitro* data and build predictive models of human disease.

ToxCast™ published the results and released data from Phase I testing of more than 300 chemicals in 2009. Moving forward into Phase II, the 700 chemicals to be tested were announced publicly later in the day. Dr. Dix mentioned that all of the publications and data related to Phase I of ToxCast™ are publicly available (*http://epa.gov/ncct/toxcast/*). He described the 320 total compounds assessed in Phase I, of which 309 were unique structures. Most (291) were pesticides, with a broad range of chemical classes represented.

He mentioned several of the partners contributing to ToxCast™ by generating HTS and genomics data including contractors, EPA labs, and Tox21 partners, particularly the NCGC. He provided details on the types of assays being conducted, both cellular and biochemical, typically numbering approximately 500. He showed data from one of the recent ToxCast™ publications (Judson *et al* in *Environmental Health Perspectives*), encompassing 467 assays used on the 320 chemicals. Dr. Dix provided other data showing the range of potencies detected for Attagene endocrine disruption (ED) activities. He also showed a spider plot depicting activity for three *peroxisome proliferator-activated receptors* (PPAR)-active chemicals.

The Deepwater Horizon disaster of 2010 represented an unexpected application of ToxCast™ capabilities, as the group was asked to screen the various candidate chemical dispersants being considered for use in cleaning up the oil spill, particularly evaluating their ED potential. Results were quickly assembled, peer reviewed, and published, facilitating an informed decision about which dispersant was best to use. The chosen dispersant showed moderate bioactivity in numerous assays, but no evidence of estrogenicity, which was the major concern.

Dr. Dix showed data depicting the fact that environmental ER active compounds span a wide potency range *in vitro*. He used those data to explain the Toxicological Prioritization Index (ToxPi), which is a visual method of depicting toxicity by weighted combinations of data, or slices, from many *in vitro* assays. A pie chart is generated, allowing a visual component to the process of prioritization. The data are organized into different domains, such as various aspects of endocrine profiling. Each domain contains information from multiple assays and multiple technologies. The ToxPi scores for all of the compounds can then be mapped to aid prioritization for further screening or testing. The ToxPi concept is continuing to be developed, with the addition of exposure information, chemical properties, and quantitative structure-activity relationship (*QSAR*) data.

Work is also progressing on developing methods for using *in vitro* assays to inform *in vivo* responses. As part of that overarching effort, the pharmacokinetics of the compounds are being characterized. One method is to use reverse dosimetry to describe the oral equivalent values for the distribution of *in vitro* AC_{50} values. This depicts the wide range of bioactivity in the compounds. By taking the bioactivity values across the different toxicity pathways, pharmacodynamics can be linked with pharmacokinetics to help understand the probability distribution for the doses that activate biological pathways. Thus, a biological pathway-activating dose can be calculated for a particular chemical for specific *in vitro* or *in vivo* targets, for particular pathways, or for particular associated human diseases or animal toxicity endpoints.

Thus, ToxCast™ information can be used to relate *in vitro* data to both human disease and to animal toxicity endpoints contained in the EPA ToxRefDB database, which houses approximately $2 billion worth of animal toxicology data. In Tox21, the EPA will also use these methods with NTP data, to bring it out in computable form in publicly accessible databases so that other researchers can also use it to identify compounds associated with different endpoints and toxicities. By combining *in vivo* and *in vitro* data, univariate associations, multivariate associations, and multicellular or systems models can be generated. That has been done already with data from rat liver histopathology from chronic bioassays from 248 ToxRefDB chemicals. Those results were combined with ToxCast™-identified genes associated with progression of rat liver lesions, generating a model depicting associations related to any lesion, pre-neoplastic lesions, and neoplastic lesions.

Dr. Dix described ToxCast™ Phase II, which will augment Phase I with 700 new, diverse chemicals. They will include pharma-donated failed drugs with pre-clinical and clinical toxicity data, which will facilitate direct *in vitro*-human toxicity comparisons. Other compounds will include 10 sponsored by L'Oreal, 50 immunotoxic chemicals sponsored by NTP, several data-

rich chemicals donated by the FDA Center for Food Safety and Applied Nutrition (CFSAN), and several "green plasticizers." He showed a chart depicting the distribution of the compounds and the availability and sources of existing data.

He concluded by describing what is ahead for ToxCast™. All data will continue to be published and made publicly available. Evaluation of new technologies will continue. Screening of endocrine disruptor activities will be accelerated. There is likely to be a Phase III (or perhaps Phase IIc), and more public meetings will be held to engender public review and broad participation in the analysis.

B. BSC Questions

Dr. Dana Loomis noted that the ToxCast™ website had a link to ExpoCast™, and asked Dr. Dix to describe ExpoCast™. Dr. Dix explained that ExpoCast™ is similar to ToxCast™, but concentrates on exposures. A database containing existing information will become available in 2011, and work is progressing on developing high throughput methods for predicting or at least estimating exposures.

Dr. Sherley asked whether it was possible to reduce the number of assays used in ToxCast™, given potential redundancies. He also asked what the plans are for developing other assays, since there is a need for additional assays to assess aspects of cell function missing from the current assay list. Dr. Dix replied that several contractors have already been dropped, because EPA didn't find their assays to be reproducible or particularly useful in terms of predictive modeling. At the same time, EPA is adding contractors and new assays. They are analyzing results from Phase I for utility and predictive modeling, but also in terms of the distribution of biological targets across different pathways and key targets within pathways. For Tox21, that is the core of the strategy for assay selection moving forward to testing 11,000 chemicals. Those are the most important decisions for the next year or so, he added. Dr. Sherley asked if there had been *a priori* discussion of the assays that would be needed. Dr. Dix replied yes, starting with the 2005 meeting and continuing with "countless" internal discussions and other public meetings.

Dr. Birnbaum asked Dr. Dix to elaborate on the issue of compound solubility, in that the compounds presently must be soluble in dimethyl sulfoxide (DMSO). Dr. Dix replied that the issue is not just solubility, but volatility as well. He suggested that one answer for compounds insoluble in DMSO might be that they are soluble in aqueous solution. He said it would be a struggle to take highly volatile compounds into the current testing paradigm. He speculated that there would be water-based assays in the future of Tox21 and ToxCast™, but that the volatile compounds would be much more challenging, as would compounds insoluble in both DMSO and water. He said a tiered testing approach using other applicable testing methodologies might address those problems.

Dr. Birnbaum asked whether the assays include some of the newer ERs, which may be reactive to chemicals that the older ERs respond to only weakly if at all. Dr. Dix replied that there is a GPR30 assay, for example, but that it had not shown great activity and had not been assessed

with a potent positive control reference compound. He said there is interest in including and expanding assays for such receptors in future Tox21 and ToxCast™ testing.

Dr. Dix confirmed to Mr. Janzen that the structures and toxicities of the failed pharmaceutical compounds would be made publicly available. Mr. Janzen asked how else the chemical industry might contribute to the program. Dr. Dix felt that this question was more appropriately addressed by those engaged in policy and regulatory activities at EPA.

Dr. Richard Miller suggested that pharma might be able to help with the extrapolation process, particularly with immunosuppressant and oncology compounds, which have extensive bodies of published literature. Dr. Dix said he appreciated the suggestions.

Dr. Justin Teeguarden wondered to what extent the EPA had tested any of the *in vitro* hazard rankings versus activity *in vivo*, or when such predictivity testing might be planned. Dr. Dix asked whether Dr. Teeguarden meant predictivity in human disease or in animals. Dr. Teeguarden replied that he meant in animals. Dr. Dix said he felt there were some strengths and weaknesses to the "animal-first" approach. He cited weaknesses in linkages between some of the endocrine activity *in vitro* and *in vivo* assays, and reproductive toxicity detected in multigenerational rat studies. The EPA has published several papers, and is publishing several additional papers describing useful new models of animal *in vivo* toxicity endpoints developed using the *in vitro* data from ToxCast™ and Tox21. However, he said, the ultimate goal is to be able to predict human disease, and environmental and ecological effects.

Dr. Lisa Minor asked about the correlation of human assay data with rat assay data from ToxCast™. Dr. Dix replied that in some cases there were data from both; there appeared to be significant conservation of some pathways, particularly nuclear receptor and signaling pathways, across species. He cited an example of a recent publication from NCCT by Imran Shah *et al* using human nuclear receptor bioactivity to stratify chemicals and predict rodent hepatocarcinogenicity. He said it is difficult to make those determinations, and so 80% or more of the ToxCast™ *in vitro* assays are based on human targets.

Dr. Bucher asked about the in-house human resources available at the NCCT. Dr. Dix replied that much of the work is conducted through contractors and the Tox21 partnership, but that the analysis and computational work are being done almost wholly in-house. He said there are 25 federal employees within the NCCT, supplemented by 15 on-site contractors or graduate or post-doctoral students. In the Computational Toxicology Research Program, which is the broader EPA activity, there are approximately 80 employees

Dr. Tim Wiltshire inquired about the range of cell lines being used. Dr. Dix answered that across the ToxCast™ assays, there were over a dozen human cell lines and primary cells being used. Work is also being done in primary rat hepatocytes. In one set of experiments, up to eight different human primary cell co-culture models are being used. Also, there are expected to be several new cell lines and primary cell assays available soon.

Dr. Birnbaum asked about the use of human hepatocytes, which are tremendously variable, as are rat hepatocytes in various strains. Dr. Dix answered that for the clearance assay,

hepatocytes from ten different donors were pooled and used. One rat strain (male Sprague Dawley) is used for hepatocytes, which he acknowledged to be imperfect. Dr. Birnbaum suggested that it might be worth considering using a pool of hepatocytes from different rats and Dr. Dix concurred. Dr. Birnbaum asked if there were any plans to use organ culture models or multiple cells. Dr. Dix said this had been done in the BioSeek cell models, and that contract proposals are presently being reviewed for complex cell culture systems. Also, he pointed to *in silico* solutions, including the Virtual Tissues Project, which involves cross-scale models of cellular organization and emergent functions, including Virtual Embryo and Virtual Liver projects.

VIII. Tox21 Partners: NIH Chemical Genomics Center

A. Presentation

Dr. Christopher Austin, NCGC Director, said he would speak about the capacities that his Center is bringing to the Tox21 initiative, rather than results on specific screens. The NCGC was founded in 2004, and currently has a staff of approximately 85 biologists, chemists, informaticians, and engineers. Its mission includes the development of chemical probes for novel biology, broad profiles of chemical libraries for biological and physicochemical properties (leading naturally to its Tox21 participation), chemical genomics (characterizing the general principles by which small molecules and their targets interact), and new technologies and paradigms for assay development, screening, informatics, and chemistry.

There is a dedicated Tox21 team within the NCGC organization, which is made possible by the existing NCGC infrastructure. All NCGC funding is external, despite its location within NHGRI. The Tox21 funding is provided by NTP and EPA. The NCGC enables the mission of all of the NIH Institutes and Centers (ICs), including Tox21, in a variety of ways.

Dr. Austin described the two screening compound collections possessed by the NCGC—the NIH Molecular Libraries Screening Center Network (MLSCN) compound collection of approximately 360,000 compounds, and an internal collection of approximately 120,000. In a given week, the Center would screen all 450,000 compounds in a 7-point dose-response, comprising between 3 million and 5 million wells per week, which is equal to the largest pharmaceutical companies. The NCGC works on the entire spectrum of biology and genomics, as opposed to the relatively small “druggable” space addressed by pharma.

NCGC has a library of more than 3,000 pharmaceutical compounds to enable drug repurposing and chemical genomics, consisting of every drug approved for human use in the United States, the United Kingdom, Canada, and several other countries. Dr. Austin described the difficulty of acquiring and curating the information on the compounds, as well as acquiring samples of the compounds themselves. The library will be included as part of the Tox21 10K phase.

He described the range of assays performed at NCGC; approximately half are biochemical, and the rest are cell-based. They include assays for phenotype, pathways and proteins.

He called the NCGC the “Grand Central Station” of NIH, which operates as a trans-NIH center, getting projects from virtually all of the NIH ICs, with the largest number of projects coming from investigators associated with or funded by the National Institute of Allergy and Infectious

Diseases (NIAID) and the National Cancer Institute (NCI). Most of the projects are with extramural investigators; currently, the center has approximately 200 active collaborations with researchers all over the world. These projects are part of the Molecular Library Program, distinct from Tox21, but they can be utilized for Tox21 as deemed useful.

Dr. Austin described the Center's quantitative HTS (qHTS). Unlike conventional HTS, which screens compounds at one concentration, qHTS assays are conducted at multiple compound concentrations, allowing robust activity profiles of screened compounds and dramatically reduced FP and FN rates. All Tox21 compounds are screened at 15 concentrations, at a range of 1 nM to 100 µM, using a 1536-well format. In this format, the plate contains 32 rows of 48 columns, the equivalent of sixteen 96-well plates. The yield after allowing for controls is 1,408 test samples per 1536 well plate. The 1536-well format is both much cheaper due to compound and reagent sparing and much faster—to test one million samples; it takes one week, compared to 4 months using a 96-well system. Dr. Austin showed the screening systems being used at the NCGC, including a new robotic screening system due to be delivered in February 2011, to be dedicated solely to Tox21 work.

With the NCGC's tremendous throughput capability generating large amounts of data, the need arises to be able to look at it in a relational way. Each of the Tox21 partners has addressed that issue. NCGC has developed the NCGC Chemical Genomics Browser as a way to visualize compounds' assay activity and dose response. The data can be grouped as desired relationally, and can be drilled into for more information (e.g., structure). All NCGC Tox21 results are made publicly available in PubChem, as well as in databases managed by NTP and EPA. One of the project's major challenges is to characterize all of the major pathways operative in mammalian cells. He said for Tox21 to be successful it is necessary to have the entire universe enumerated, allowing the development of the minimal number of assays necessary to cover the entire pathway space. The presentation by Dr. Ruili Huang later in the BSC meeting presents NCGC's progress towards this goal.

Dr. Austin concluded his presentation by summarizing the value that NCGC brings to the Tox21 MOU: (1) unparalleled screening technologies and production pipeline with a unique qHTS paradigm developed originally for probe discovery but ideal for Tox21; (2) highly experienced scientific staff from the best pharmaceutical, biotechnology, and academic organizations; (3) expertise in assay development, optimization, cheminformatics, follow-up assays, and analytical and synthetic chemistry; (4) through the ML and other NCGC programs, availability of a very broad range of assays in virtually every area of biology and disease; and (5) experience and focus on "big science," with a collaborative, team-based, and deliverable-based culture.

B. BSC Questions

Dr. Birnbaum asked Dr. Austin how many assays the NCGC is routinely running. He replied that assays are typically run in series rather than in parallel, at least within Tox21. For Tox21, he estimated that the NCGC runs about one assay per week, or 40-50 assays per year. He noted that different assays yield different readouts, ranging from just one to up to 100. Dr. Birnbaum inquired about the perception that it is difficult to get a chemical of interest added to

the collection. Dr. Austin replied that it had been urged that the collection be as big as possible from its inception, so that it would be comprehensive from the beginning. However, he said, there are often smaller runs of compounds for particular reasons, citing the Deepwater Horizon work as an example.

Mr. Janzen asked Dr. Austin to confirm his impression that the NCGC does confirmations on all active findings. Dr. Austin replied that at the beginning they had done so, but that currently they do not routinely do so, because the goal is for the profiling data to stand on its own as much as possible. By retesting any of the compounds, it would quickly become a situation where everything would need to be retested—or nothing. As it is, in the Tox21 Phase II library, every compound will be tested three times at 15 concentrations in every assay, essentially doing the confirmation assay up front, not once, but twice.

Mr. Janzen asked whether the new robotics system coming to the NCGC would only be capable of 1536-well runs. Dr. Austin replied that it would be capable of multiple formats, including 384- and 96-well, because not every assay is appropriate for the 1536-well format, one example being the micronucleus assay. He estimated that 90% of the assays would be 1536-well.

Dr. Sherley asked whether there is sufficient depth in cell assays to start asking about the time vector of activity. Dr. Austin said that before any assay is run, part of the validation/optimization process is to run a time course assessment to determine the time point that will give the greatest sensitivity. He said what has been seen is that some compounds act early, some late, and some exhibit bell-shaped activity. Generally, a single time point that seems to yield the greatest sensitivity is chosen.

Dr. Novak asked how NCGC handles issues related to cell-based incubation such as confluency or autocrine effects. Dr. Austin replied that those issues are addressed in the validation process. Generally, a level of confluence that gives the highest sensitivity is chosen, since FPs can be dealt with but FNs present a thornier problem. He said the possibilities for testing are virtually infinite given the many variables, but that the team must focus its efforts on pathways it believes to have predictive value.

Noting that NCGC puts all of its data into the public domain, Dr. Bucher asked how much time the organization has to analyze and publish. Dr. Austin replied that what happens to the data depends completely on the funder. In Tox21, the rule is that once the data are published, they are deposited into PubChem. That is different from the rule followed in the ML program, which mandates that data must be released within two weeks of validation. Dr. Birnbaum asked why the rules are different. Dr. Austin explained that the practice with Tox21 is more typical, and that there is much more of a competitive atmosphere surrounding the ML data. With Tox21, he said, it is still an evolving question, with the sense being that they want to have confidence in the data and ascribe context and meaning to it before it is made public, due to the high potential for misinterpretation. Dr. Tice added that care has been taken with the Phase I data to ensure that the assays and data outcome were understood. He said the paradigm will shift with the 10K library, because the data cannot be held up for up to two years—it will be more like the ML initiative in that once the data are generated they will be released. He noted that the three

partners release data to three databases—NCGC to PubChem, EPA to Aggregated Computational Toxicology Resource (ACToR), and NTP to Chemical Effects in Biological Systems (CEBS). Dr. Birnbaum was pleased to hear that Phase II data will be released quickly, noting Dr. Collins' mandate that all data be released rapidly.

IX. Tox21 Partners: U.S. Food and Drug Administration

A. Presentation

Dr. R. Daniel Benz, U.S. FDA, described the FDA's contribution to Tox21 by first outlining resources available to the Tox21 partners from the FDA's National Center for Toxicological Research (NCTR), including (1) the Liver Toxicity Knowledge Base, which contains information on more than 1000 drugs with data related to drug-induced liver injury; (2) the NCTR Liver Carcinogenicity Database, which contains 999 chemicals with liver carcinogenicity data; and (3) the Endocrine Disruptor Knowledge Base, which contains data on more than 5,000 endocrine-active chemicals and controls. He noted that NCTR would also provide advice concerning methods for bioinformatics approaches to analyze large data sets, and would participate in data analysis as appropriate.

Dr. Benz described the unprecedented donation of data on failed drugs by the pharmaceutical industry, which has emerged with help from the FDA Center for Drug Evaluation and Research (CDER), Office of New Drugs. He also noted that although FDA cannot legally share specific human data on specific drugs, there is no restriction on sharing general scientific knowledge derived from them, which may constitute useful information.

Dr. Benz mentioned that he works with the CDER Office of Testing and Research (OTR), which includes the QSAR Computational Toxicology Group. It is an applied regulatory research group that (among other activities) provides computational toxicology evaluations for drugs, metabolites, contaminants, excipients, degradants, etc. to FDA/CDER safety reviewers. The group's scope is to predict accurately, with *in silico* software, *in vitro*, animal and human effect endpoints of interest to CDER. Its goal is to speed development of safer drugs by early identification and elimination of safety concerns for active pharmaceutical ingredients, metabolites, and impurities.

Dr. Benz pointed out that for Tox21, appropriate models of the group's 105 *in silico* QSAR models can be used to predict toxicity of diverse chemicals, using a range of approaches to arrive at predictions using several (Q)SAR computational toxicology software programs. He said the group also has a database of approximately 16,000 drugs or chemicals associated with drugs (e.g., excipients). Detailed toxicology study information is also available in XML format through Leadscope, one of the group's collaborators. The group has agreements with five companies (including Leadscope) that have prediction software operating in distinctive ways.

Dr. Benz delineated the various non-clinical effects models used, including six models of rodent carcinogenicity and 11 models of genetic toxicity. He also described the group's human clinical adverse effects models, using human data to predict human endpoints. These prediction suites, which are based on data points consisting of adverse event reports to the FDA, include

hepatobiliary (5 models), renal/bladder (6 models), cardiological (13 models), pulmonary (21 models), and immunological (19 models). He said some of the models within the last two groups might not survive an upcoming validation exercise. Other organ systems will be included in the future. There also are two models for maximum recommended daily dose estimation.

He predicted that QSAR computational toxicology would be part of a new safety assessment battery, along with HTS, various flavors of –omics, mode of action analysis, and additional newly developed methodologies. However, he speculated that there would still be some *in vitro* testing, rare and highly specific animal testing, and reduced human clinical testing in the future.

B. BSC Questions

Dr. Birnbaum said Tox21 is "really thrilled to have such active FDA participation," and noted that CFSAN is involved. Dr. Howard said interest in Tox21 at FDA is "percolating," as evidenced by the number of FDA members in the working groups.

Dr. Teeguarden asked Dr. Benz to elaborate on the validation of models, the standards for validation, and how well the models have to function to be useful. Dr. Benz replied that the FDA uses two external validation methods and the standards up to now have been 85% specificity and 90% coverage. He said the sensitivity sometimes suffers in proportion to the size of the training set. He said the models are used in five different platforms, with any one being positive being interpreted as a positive overall call.

Dr. Sherley asked about the apparent barrier regarding human data, wondering whether it matters where the activity takes place with regard to HTS predicting human health effects of compounds, in that HTS data could be moved into the FDA for predictive modeling computation. Dr. Benz felt that it certainly might be hypothetically do-able if the personnel would be available to perform the analyses. Dr. Tice said it had been discussed to send the FDA the structures of the 10,000 Phase II Tox21 compounds to have them run them through their models. It can be done, but will take time, as they are not set up to run 10,000 compounds in a single run. He said Tox21 wants to be careful not to use "black box" prediction models, so there has been some reluctance to embrace the models, since they cannot be externally validated. He said there is a Small Business Innovation Research (SBIR) project in process with Leadscope to take all of the FDA data and the current Tox21 data, and link it all together. That is the ultimate goal, but tools to accomplish it are still in development. Dr. Benz clarified that all of the software and models used at FDA/CDER are available to the general public, but must be purchased.

C. BSC Discussion

Dr. Solomon, first lead reviewer, said the presentations were impressive and that it is exciting to see how far the program has moved in relatively little time and with relatively slim resources. She said the goals of the program are quite clear and remain relevant to public health. She recommended reassessment of the tools being deployed to ensure that they are optimal to meet the program's goals, which is "a moving target," given the rapid development of the science and improvement of the available tools. She wondered whether some of the more high-content

assays might be useful to pursue at this point, particularly as some of the genomic assays have become cheaper and more available. She noted that some of that is already happening.

Regarding the adequacy of the collective scientific and technical capabilities of the participating agencies to achieve the goals of Tox21, she felt that exposure and human epidemiology were perhaps underemphasized. She said there were some assays that could be used not only to screen chemicals, but also to screen human cohorts, such as occupational workers. She wondered whether it might be appropriate to include some portion of the Centers for Disease Control, either National Institute of Occupational Safety and Health (NIOSH) or the National Center for Environmental Health, or both, in Tox21.

As to suggesting areas where increased scientific emphasis or resources could be most beneficial, Dr. Solomon recommended more emphasis on multiplex assays, more emphasis on human data or epidemiologic studies, and starting to put together the exposure piece of the puzzle, which she considered extremely important. She also found efforts to analyze HTS data to be quite important, including critical assessment to ensure that the information being generated is the most useful for informing decisions. Gap analysis should be continued, she added.

Dr. Miguel Fernández, second lead reviewer, said he comes from a clinical background, and so would focus his comments on aspects that relate to clinical matters. He agreed with Dr. Solomon's points and said the clarity of the program is appropriate. He considered the public health relevance an area that may need more attention, in terms of a continuing effort to educate the public, so that people understand what is being done in the program and its relevance. He said it would be critical to make such efforts, particularly so as to ensure funding in the future. Regarding the adequacy of the scope of the Tox21 program, he was concerned about the difficulty of assaying volatile chemicals, adding that it would be important to address that, because volatile chemicals are often in pesticides and are involved in worker exposures. He worried that "some real, applicable science" was being missed and asked what the next step would be once the several hundred thousand chemicals had been screened. Dr. Fernández suggested looking at the interaction between the human genome and the microbiome, and that not doing so would be missing important information about how toxicants and toxins interact with human cells. The next dimension would be to include all of the host organisms, as opposed to dealing solely with human cells in isolation.

Dr. Loomis concurred with the other reviewers regarding the great promise of the Tox21 approach. He agreed with Dr. Solomon that the need for human data had been underemphasized; human epidemiology was the only way to provide directly relevant data from the right species. He said it was unclear from the presentations how that type of research was going to be supported and integrated into the program. He also recommended more integration of human exposure assessment, which is vital to assessing risk at the end of the screening process, and for setting priorities regarding which compounds to look at in the first place. Dr. Loomis was impressed with the presentations regarding the Tox21 partners and the capabilities described and noted that two of the three major areas covered by regulation, food, drugs and cosmetics, and the environment, were represented, but that the third, the work environment,

was not. He questioned whether NIOSH was in the process, that despite the fact they don't do HTS, NIOSH would be a valuable source of human exposure and epidemiological data. Regarding resources, he reiterated the need to integrate human exposure data and health outcomes. He lauded the fact that "at last" the capability exists to evaluate large numbers of compounds at lower cost and at greater speed than ever before. Although human exposure and epidemiological studies are slow and expensive, they still need to be done, and resources should be directed toward those endeavors.

Dr. Tice responded by stating that the Tox21 projects linkage to humans would become clearer in subsequent presentations about the Tox21 working groups and activities. He agreed that the linkage to humans would be critical to determining whether or not there are actually prediction models for human adverse outcomes or human disease. The FDA would be able to provide some of the vital information needed, including its models. The need to link animal study data to humans is also recognized. He said metabolomics was being examined as a tool for doing so; one example for relating human and animal metabolomics data might be with herbal products.

Dr. Tice acknowledged the need for exposure information to inform prioritization. He said the EPA is active in that area (e.g., ExpoCast™), as is NIEHS, which is developing exposure models in human populations. Also, he mentioned that NTP supports the Comparative Toxicogenomics Database, a publicly accessible database that has been funded to bring exposure information into the database and into its genomic analysis.

He agreed that high-content screening might potentially be a second level of assays for Tox21, particularly as a way to get more information about compounds emerging from the 10K library. He described several ongoing efforts in that area, including companies with functional 3D organ models.

In terms of public education, he felt that the Tox21 partners have been succeeding at interacting with the scientific community, with almost 200 presentations at scientific meetings or to interested scientific organizations since the MOU was signed in 2008, but that interaction with the non-scientific public needs improvement.

He said the issue of volatiles is clearly a limiting issue for *in vitro* assays, but that the focus has been on characterizing pathways prior to exposing them to something that is more difficult to measure. He said it is an issue on the list to be considered.

Regarding organism interactions, he mentioned the NIH Microbiome Project, and said Tox21 would await results from it before determining how to integrate that element. He added that there is work going on with *C. elegans* and zebrafish, for more complex organism information.

For human epidemiology data, he mentioned the upcoming NIEHS workshop on obesity and diabetes. Some of the assays in Tox21 were chosen based on working backwards from epidemiological data. The recent NIEHS workshop on the genetics of autism was another example of using that approach. He also cited an ongoing epidemiological study involving the incidence of uterine fibroids in a population of African American women in western North

Carolina. The speculation is that exposure to skin lotions may be a contributing factor, so the lotions are being tested for estrogenic activity. Regarding human exposure assessment, Dr. Tice pointed out that NIOSH is part of NTP, and so is in fact involved in Tox21.

Dr. Bucher noted that the comments regarding human exposures and the need to link assay results with human outcomes bring up the same issues faced by all of NTP's assays, not just the HTS work, although the HTS accelerates the need to confront those issues. He said there are interagency agreements in place with NIOSH for exposure information that could be exploited in the future. There have been numerous associations noted in epidemiology studies, but the HTS data allow them to be approached from a different direction, helping establish biological plausibility for some of the associations. He noted that the upcoming workshop on diabetes and obesity would be a good test case for developing that approach. There are plans to do additional workshops on cardiovascular and neurological diseases soon. Thus, there will be the opportunity to integrate information from HTS with information from traditional toxicology databases and with information from experts in the various diseases, who can advise on pathways and targets for further HTS studies, facilitating the next generation of such assays.

Dr. Birnbaum noted the cross-NIH effort looking at the microbiome, the meeting held two weeks earlier at NIEHS regarding the microbiome, and an upcoming Emerging Issues meeting on the microbiome to be held at the National Academies in April 2011. She said the NIEHS is currently co-funding a National Research Council (NRC) effort to develop exposure assessment in the 21st century, a parallel effort to Tox21. She added that human pharmacokinetic research studies are being conducted on agents of high interest at the NIEHS clinical research facility, such as a study of cashiers and their exposure to BPA from handling register tapes. She said she liked the discussion about educating the public on these issues, and that with limited budgets expected in the near future, it will be important to communicate more and more about the concept of predictive toxicology. She suggested starting with an article for the informed public in *Scientific American*, which would be likely to be circulated in the mass media.

X. Tox21 Working Groups: Introduction

A. Presentation

Dr. Tice provided a brief introduction to the Tox21 Working Groups (WGs). He depicted the organizational structure of Tox21, including the four WGs: Chemical Selection, Assays & Pathways, Informatics, and Targeted Testing. Each of the WGs is co-chaired by one representative from each of the four Tox21 partners, along with other members. He showed a graphical representation of the interactions among the groups and their relationships to the NAS report concepts about toxicology in the 21st century. He described the upcoming afternoon poster session for the BSC members, briefly summarizing the list of posters and their relationship to the different sets of presentations.

XI. Tox21 Working Groups: Chemical Selection Working Group

A. Presentation

Co-chair Dr. Cynthia Smith, NIEHS/NTP, briefed the BSC on the Chemical Selection Working Group (CSWG). She noted that having a large, diverse, well-supported chemical library would be a cornerstone of the Tox21 initiative activities, and that construction of the library falls to the CSWG. The group originated with the Tox21 MOU, and its purpose is to (1) coordinate the selection of compounds, (2) standardize preparation of plates for inclusion in the Tox21 library, (3) develop a documentation scheme for compounds included in the library that tracks with assay data, (4) devise a quality control (QC) approach to support compound data, and (5) expand beyond current technical limitations.

In the first NTP phase of HTS, which started in 2005, an initial library of 1408 compounds was developed, with compound selection focused on those that had been tested in one or more standard NTP assays, along with a few other sources. Dr. Smith said there was a significant learning curve for NTP at that point in terms of understanding the many requirements of HTS, and speculated that EPA had had a similar experience with Phase I of ToxCast™. Those lessons learned were valuable in efforts to develop the much larger 10,000 compound library planned for use at the NCGC in Tox21 Phase II. The plan was to develop a 10K compound library designed to maximize coverage of chemical space, including all possible chemical classes, with QC built in to support data use. To start the effort, lists of potential compounds were compiled from published sources in the ACToR database, internally invited compound lists, and other submitted lists. These lists were eventually pared down to an "HTS-able" working list.

To develop the "HTS-able" library, the initial list of approximately 120,000 compounds was first screened for duplicate or replicate chemicals, which yielded about 19,000 unique compounds. Other screening mechanisms ultimately yielded a final "HTS-able" list of about 11,000 compounds. The effort to obtain these compounds was then divided among the (then) three agencies involved. The NCGC took responsibility for drugs, drug-like compounds, and active pharmaceutical ingredients as they had many of these on hand. The EPA took responsibility for the ToxCast™ I and II compounds, compounds from the Antimicrobial Registration Program and the Endocrine Disruptor Screening Program, the OECD Molecular Screening Working Group list, failed drugs being provided by different pharmaceutical companies, and other compound lists of interest to the Agency. The NTP took responsibility for NTP-studied compounds of all types, NTP nominated and related compounds, ICCVAM and NICEATM validation and reference compounds, and compounds from outside collaborators such as the U.S. Army Public Health Command.

It was determined that it would be desirable to have about 100 intentional duplicates on each 1536-well plate used for screening in qHTS at the NCGC. These compounds were identified by prioritizing the Tox21 Phase I data, filtering the resulting candidates according to availability and physical/chemical properties, with a focus on those that behaved best in solution. This process ultimately resulted in the selection of 88 compounds to be used as internal reference compounds. These were then procured and formulated under an EPA contract, and the compounds were distributed to NTP and NCGC on 96-well plates.

The NTP conducts QC analysis to confirm the identity and purity of the bulk materials before the DMSO solutions are formulated. This confirms the information from the suppliers' Certificate of Analysis. The process, involving mass spectrometry, is then also valuable as an independent quality check on a portion of the compound library.

Dr. Smith showed a graphic depicting the chemical space covered in the library, including the compounds being screened by the various partners. She said a few of the compounds shown on the graph would probably be removed from the list once the chemicals were plated.

She described the plating strategy utilized by the group to arrive at uniform 1536-well plates. Ultimately, each 1536-well plate will contain two of the sets of 88 designated duplicates, randomly distributed on the top and bottom of the plate. NCGC, which will assemble the 1536-well plates, will be sent 9 copies of each of the 384-well plates, along with one set to be sent out for QC analysis. Several steps will take place at NCGC, with compound and plate IDs verified and preserved at each step. Eventually, six 1536-well plates will be generated from each set of 384-well plates. One plate at a time will be used for assays, with the others stored at -80°C. The plates are used at room temperature, with a shelf life of about five months, yielding 2-1/2 years of usage of the six plates.

Dr. Smith said 100% analysis of library compounds is a requirement. Tox21 has contracted with a company with considerable experience in the process with pharmaceutical HTS. It uses an analytical method that has a 3.5-minute run time per well, which is not extremely rapid but is still preferable to one-at-a-time laboratory procedures. The platform is based on liquid chromatography with mass spectrometry (LC/MS), using positive and negative electrospray ionization (+/-ESI), but also with evaporative light scattering detection (ELSD), chemiluminescent nitrogen detection (CLND), and ultraviolet diode array detection (UV-DAD) in line with the LC/MS. The process yields confirmation of identity and purity, and for some compounds, concentration. Compounds will be re-analyzed for stability after a period of use. It is understood that this process will not work for all compounds, so a subset of compounds will be QC'ed using another method.

The Tox21 Compound Registry will be maintained by EPA, using its DSSTox schema for structural and compound information along with unique solution IDs for each compound issued by Tox21. All files will link to the QC analysis data.

Summarizing the status of the library, Dr. Smith said the NCGC has already prepared two plates, one of which has undergone QC, with the other currently in the QC process. She said a process is being developed to grade the plates that emerge from QC according to the compounds' ultimate quality and usability. The EPA has just issued an order for preparing its first plate set. The NTP has just received duplicate compounds and is preparing to plate its first set. The Tox21 10K library (which is actually more than 10,000 compounds) should be finalized by February 2011. The NCGC expects to bring its new robotics station on line by spring 2011, and by summer of 2011 QC should be complete.

Dr. Smith mentioned some of the ongoing collaborative efforts that are underway, including work with the NIH ML Small Molecule Repository, and with extramural scientists who are chiefly

working on SAR models or on specific assays. The working group would like to tackle three main issues in the future: solubility, mixtures, and QC at the assay level.

She said the CSWG has worked well together as a cooperative effort among chemists at the participating agencies. It has developed a plan for at least six 1536-well plates, has devised and applied a QC plan for the library, has developed a compound registry to accompany assay data, and plans to continue to work toward increasing the size of the library.

B. BSC Questions

Dr. Birnbaum noted that there are currently some 80,000 chemicals in commerce, with very little toxicity data on any of them. Given the CSWG experience in selection, with the large number of duplicates detected, she wondered if the same phenomenon might be true of the commercial chemicals. Dr. Smith said she felt that is probably true, including duplicates and different grades of the same chemicals. Dr. Birnbaum asked if the working group had looked at the different stereoisomers of the compounds. Dr. Smith replied that there were a few cases where stereoisomers had been included if the different isomers were of interest, and that they would be put in separate wells.

Dr. Sherley asked for more detail about the inability to know the quantity of a compound in a final assay. Dr. Smith said it addresses the question of what happens to the compound in the buffer, and whether negative responses in the assays may be in part due to the compound not really getting to the cells. Stability is also a question. Dr. Sherley asked how much of a problem these uncertainties are. Dr. Smith said they certainly need attention, and an approach needs to be developed, but she was unable to attach a particular number in terms of the urgency of the issue. Mr. Janzen considered this a “huge issue,” citing a paper he had published. Dr. Smith requested suggestions for addressing the issue.

Mr. Janzen asked about the diversity of the chemicals chosen. Dr. Smith confirmed that they had been chosen initially if they appeared on a list, or on lists sent by outside parties, such as lists on EDs, fire retardants, and others. Mr. Janzen asked about the group’s goal of 90% purity. Dr. Smith replied that the goal was not only diversity in the library, but that purity compared to solution QC was important as well. Mr. Janzen asked about the compilation of the sets by the Tox21 partners. Dr. Smith elaborated on that process, and mentioned that there remains some overlap in the lists that needs to be resolved.

Addressing the QC issue, Dr. Carney asked if there is a process to review the data and check on anomalous results. Dr. Smith said the CSWG will be capable of doing so, but had not done so yet. She said Tox21 expects the HTS QC approach to cover at least 70% of the compounds, and that single compound studies will be conducted after the HTS QC runs are complete, as necessary. Compounds with purity less than 90% will be flagged accordingly.

Dr. Birnbaum asked whether solubility could be enhanced by the use of carrier proteins, such as albumen. Dr. Smith said that idea had not been addressed, but would need to be eventually, as would the issue of volatility.

XII. Tox21 Working Groups: Assays and Pathways Working Group

A. Presentation

Ms. Kristine Witt, NIEHS/NTP, briefed the BSC on the Assays and Pathways Working Group (APWG). She said members of the group participate in biweekly teleconferences to review assay nominations and conduct other business related to the APWG's responsibilities. Those responsibilities include: (1) identifying targets for screening, such as enzymes, gene regulation, pathway activation, and biochemical reactions—targets should have biological significance to toxicological endpoints; (2) identifying appropriate assays to screen the desired targets; (3) tracking screening activities, reviewing assay results, evaluating assay performance, and troubleshooting assay problems; and (4) conducting outreach to stimulate assay design and acquisition, and to promote the research and development of new assays and technologies.

She reviewed the APWG goals in Phase I of Tox21, noting that the experience of the past five years would be vital to understanding how to move forward with Phase II. First, she said the APWG needed to determine which targets would be most useful to screen in terms of providing information necessary for toxicological profiling, and then prioritization for follow-up testing. Second, the APWG conducted a thorough evaluation of a variety of assay formats, and evaluated cell types such as established cell lines and primary cells, and investigated interspecies differences in cellular response. Also, the APWG needed to decide whether to concentrate on human cells or rodent cells. The group also needed to define criteria for acceptable assay performance, both in technical and in biological terms. And finally, the APWG needed to assess whether HTS, a technology developed for drug discovery, could be applied to environmental compounds to measure potential toxicity.

The APWG also considered what would be appropriate endpoints for measuring xenobiotic exposures. They considered all types of toxicity, such as carcinogenicity, neurotoxicity, and genetic toxicity. They considered the merits and drawbacks of selecting targets at the beginning of a signaling pathway, at network nodes where two or more pathways might intersect, or at the end of a complex signaling cascade. The APWG also considered whether any and all indicators of cellular perturbation are potentially useful measures of toxicity.

Ms. Witt reviewed the qHTS protocol that had been developed, which she described as an extremely critical development in applying HTS to toxicology. The assays used in the protocol have specific requirements, which affect the process of assay selection. For example, the use of the 1536-well format reduces the volume per well to approximately 5 µL, with no aspiration steps, thus requiring all assays to be homogeneous. Due to the small volume per well, total incubation time in an assay is usually held to under 48 hours to avoid confounding by evaporation. With the reduced number of cells per well, assays need to provide robust response signals for detection. Assays should be pre-validated in 96- or 384-well plates and performance must be acceptable, with a signal-to-background ratio >3, a coefficient of variation <10%, and a Z factor >0.5.

Early in Phase I, extensive proof-of-principle studies were conducted to confirm the applicability of qHTS to environmental compound screening, using commercially available, well-

characterized assays. Five different endpoints were tested, several potential confounders were examined, and numerous cell types were tested. Following proof-of-principle, assay identification and acquisition proceeded through a number of avenues, including a workshop held at NIEHS in September 2008, outreach to the international community and the NIEHS and EPA scientific communities, as well the circulation of an assay nomination form. Several assays were acquired through each of those mechanisms.

As assay acquisition was taking place, the APWG established a screening workflow, beginning with the nomination of assays from a variety of sources, through an approval or disapproval process, followed by screening of acceptable assays, to the ultimate result of public release of data at the culmination of the process. Ms. Witt described the 102 qHTS assays that were screened at NCGC during Phase I. They included assays related to apoptosis, cell viability, DNA damage, epigenetics, mitochondrial toxicity, nuclear receptors (NRs), and stress response pathways.

Ms. Witt said the APWG had met its goals for Tox21 Phase I, having (1) identified a variety of endpoints related to toxicological responses, (2) evaluated a variety of assay formats, (3) evaluated cell types from several species, with an emphasis on human cells, (4) defined technical and biological criteria for acceptable assay performance, and (5) demonstrated that HTS could be used for toxicity profiling of a library of environmental compounds. The group also developed a strategy for determining mechanism of action, moving from the information emerging from a primary screen through a series of increasingly focused follow-up assays, arriving at a smaller group of assays designed to confirm the primary screen in terms of mechanism of action.

The APWG has delineated a Phase II qHTS screening strategy. It is based on Phase I experience, discussions among the pertinent parties, information from *in vivo* toxicological investigations, maps of disease-associated cellular pathways, and recommendations from the July, 2010 Assay Selection Strategy meeting. It has been proposed that the first stage of Phase II strategy focus on induction of stress response pathways and NR activation or inhibition.

Stress response pathways screening was chosen because they are protective signaling pathways activated in response to environmental insults. They are highly conserved, broad indicators of early cellular perturbation. They are triggered at low doses before other effects (e.g., cell death, apoptosis) occur and their mechanisms are well characterized. The transcription factor/sensor complex integrates multiple signaling inputs. And finally, they were preferred because patterns of activation vary by compound, and so a battery of such assays can be used to build compound-specific stress response profiles. Ms. Witt showed a chart depicting several of the major known stress response pathways, their inducers, and the transcription factors and genes involved. They included pathways of oxidative stress, genotoxic stress, heat shock, endoplasmic reticulum stress, hypoxia, inflammation, and metal response.

She elaborated on the reasons for screening for NR activity: (1) NRs are sensors of small molecules that regulate gene expression controlling development, homeostasis, metabolism, and detoxification pathways; (2) ligand binding domains are lipophilic pockets evolved to bind

either high affinity or moderate-low affinity endogenous ligands; (3) xenobiotics are often lipophilic small molecules; and (4) xenobiotics such bisphenol, genistein and dichlorodiphenyldichloroethylene have the potential to interfere with important signaling pathways.

There are 48 human NRs, and participants of the July 2010 Assay Selection Strategy meeting advised screening all of them, if appropriate assays could be identified. Ms. Witt presented a list of approximately 52 assays the group plans to screen initially in Phase II, including assays for stress response pathways and NR activity.

As assay data and assay selection strategies are evaluated, several critical issues related to qHTS data interpretation and application must be kept in mind. For example, there is currently no method to introduce xenobiotic metabolism. Also, qHTS does not consider interactions between chemicals or different cell types. It remains to be determined how to screen for chronic exposure-induced toxicity, as well as how to accurately extrapolate from *in vitro* concentration to *in vivo* dose.

The APWG plans to develop an approach using bioinformatics that considers disease from multiple levels, including pathways, biological process networks, intermediate phenotypes, and tissues, organs, and cell types, as well as genes. The plan is to develop assays that will query disease-specific health effects resulting from xenobiotic exposures. Ms. Witt illustrated the concept with a flow chart ranging from assay to disease. Identification of disease-specific assays will be a major focus of the working group in the coming year. She concluded her presentation by mentioning that one of the methods of developing new assays is through NIEHS Small Business Innovation Research/Small Business Technology Transfer (SBIR/STTR) contracts, and she listed several that have been awarded during 2010 and announced for 2011.

B. BSC Questions

Dr. Birnbaum asked about the use of full-length and ligand-binding domain (LBD) assays in related NRs. Ms. Witt replied that there had been much discussion of that issue within the working group and in Tox21 regarding the benefits or drawbacks of using either approach, with two distinct schools of thought. Thus, the decision was made to run both, on both the ER and androgen receptors (AR), and use the data to resolve the questions involved. Dr. Birnbaum cautioned that using only the LBD would result in missing information. Ms. Witt said the group is working to identify sources for full-length assays.

Dr. Carney asked about proof-of-principle, noting that the group seemed to have been concentrating on the mechanics of the screening, rather than seeing whether the screening is actually predictive in terms of adverse outcomes and useful in risk assessment. He wondered whether there was a plan to take some model pathways and run them all the way through the system. Ms. Witt said that was an approach that had not been used yet, although it was a good one. She said they had taken a similar approach in terms of characterizing mechanism of action, using follow-up studies to confirm whether they were going in the right direction in making biological sense. She said in every case, the follow-up studies showed that the primary screen had been detecting the activity they thought they had been detecting. So there is

confidence in the screens, but she concurred with Dr. Carney's comments. He said he felt that the more comprehensive approach was probably what the public would look for. Dr. Tice mentioned that one aspect being studied in Tox21 is the endocrine disrupting compounds, including well-characterized and less well-characterized compounds. He said that the Tox21 data are being compared to existing *in vitro* data for the compounds and their effects on ER and AR with regard to potency and efficacy. In response to Dr. Carney's proposal, he said there are certain areas that are easier than others to go from "top to bottom," while there are other areas where it's more difficult to do so. He reiterated that Phase I was technically oriented, looking for proof of principle that the studies could be carried out and useful data could be generated. It was necessary to be confident in the project's ability to clearly classify actives prior to moving forward into studies such as those suggested by Dr. Carney, as well as having easy access to multiple sources of relevant data, such as CEBS. He said now that Tox21 is reasonably confident in its analysis tools and resources, by the next review, there should be an "explosion" of the kinds of studies being discussed.

Dr. Sherley asked for more clarification regarding assay selection, in terms of the overall cell space and the chemical space, and how many processes in the cell can be, or need to be assayed. He also asked about the list of needs for new assays. Ms. Witt replied that there is a long list of needed assays. Assays in the first stage of Phase II screening will measure early events triggered by xenobiotic exposure. She said the group is in communication with a number of different vendors and assay developers, as well as NIEHS intramural scientists and EPA personnel who could develop assays to meet specific needs. Dr. Sherley asked what the scope of the problem is. Ms. Witt estimated that at least 100 or more assays were needed. Dr. Tice added that the goal for the screening taking place at NCGC would be to find assays that cover broad biological space, integrating multiple pathways. Some of the compounds might be picked out for inclusion in ToxCast™, which has the ability to run about 500 assays in a relatively short period of time, compared to the much smaller number of assays utilized in screening at NCGC. As more information becomes available on which assays are necessary to characterize toxicity and which may not be additionally informative, it is anticipated that the number of assays actually needed would be reduced. At this time, he said, the number of assays required to generate complete toxicity profiles on environmental compounds is just not known and really can only be estimated.

Dr. Zelikoff asked about the cell lines, and whether there had been any attention to specific organs or organ systems, in that different chemicals would have different mechanisms depending on the organ or organ system it might be involved with. Ms. Witt said that early in the process, the original 13 cell lines screened were chosen because they came from organs that were typical sites of toxicity resulting from xenobiotic exposures. She said there is a trade-off in terms of resources and effort between screening one endpoint across multiple cell types versus screening multiple endpoints in one or a few cell types. The proposed strategy is based on generating data from more endpoints in fewer cell types in order to produce a broad range of data necessary to prioritize compounds for additional follow-up studies. She said in the initial stage of Phase II, a number of different cell types would not be investigated for each endpoint, keeping the number of cell types small to facilitate comparisons of responses across assays. Dr. Zelikoff said she understood the issues, but hoped that the cell lines would be selected

keeping in mind the goal of evaluating human health effects. Ms. Witt stated that the group was aware of the importance of cell type to the biological relevance of a particular response.

Dr. Fernández asked about the cell lines and screening methodologies in terms of acute versus chronic toxicity, and whether there is an ability to distinguish between the two. Ms. Witt replied that it was unknown whether the current screening strategy using qHTS and short-term assays provided information pertinent to chronic exposures and related toxicities. Dr. Fernández pointed out that in different cell types, the effects of chronic exposure might be quite different from acute exposure. Ms. Witt agreed.

XIII. Tox21 Working Groups: Informatics Working Group

A. Presentation – Overview

Dr. Keith Shockley, NIEHS/NTP, one of the co-chairs, presented an overview of the Tox21 Informatics Working Group (IWG) to the BSC. After reviewing the Tox21 organizational chart and the group's membership, he described the working group as being about how to answer different questions arising from different data structures. He said it's an iterative process; iterative between the different members of the working group, who come from different backgrounds in quantitative fields, and as they interact with other toxicological experts.

He outlined the IWG's goals: (1) to develop computational procedures for distinguishing between active, inactive, and inconclusive responses; (2) to develop informatics tools for evaluating the results obtained from testing conducted in support of Tox21 for predictive toxicity patterns; and (3) to make all Tox21 data publicly available, to encourage independent evaluations and/or analyses of Tox21 test results.

Dr. Shockley reviewed the Tox21 Compound Libraries, which generate the data structures for the IWG's efforts. They include the Tox21 Phase I libraries, the NTP-1408, and the EPA-1408 along with the EPA-54. ToxCast™ Phase I consists of 320 compounds. Tox21 Phase II will be more than 10,000 compounds; ToxCast™ Phase II will be at least 700. He also related the various concentrations and duplicates involved with each phase. He detailed the various data processing methods being employed by the IWG, including (sequentially) methods of normalization, outlier removal, and flagging compounds.

Dr. Shockley described how activity calls are made from normalized qHTS data, which is one of the central questions for Tox21. He showed a graph of data points depicting concentration and response, and described the mathematical processing required to make an activity call, i.e., whether the data show that a compound is active, inactive, or inconclusive. The different organizations use different algorithms to suit different purposes, including (1) assay-specific analyses for ToxCast™ and Tox21 Phase I qHTS (EPA), (2) curve class for Tox21 Phase I qHTS (NCGC), (3) decision tree for Tox21 Phase I qHTS (NTP), (4) mathematical modeling for Tox21 Phase I qHTS (NTP), (5) preliminary Test Estimation for Tox21 Phase I qHTS (NIEHS), and (6) other methods.

He said the NCGC curve class approach has been revised from its original pharmaceutical orientation to make it more suitable for toxicological efforts. The NTP decision tree algorithm is

used because toxicology must answer multiple questions in order to make an activity call. He showed several examples of activity calls, including how they each scored in the NTP and NCGC algorithms.

Dr. Shockley described methods of chemical prioritization for toxicity testing being employed in Tox21. They include staged evaluations such as rank ordering, pair wise comparisons, WormTox modeling, and ToxPi, as well as prediction modeling methods such as the NCGC BioPlanet of Pathways, weighted feature significance models, and other methods. To illustrate the concept, he shared an example of data from a pair wise comparison between two profiles.

He also showed examples of how the group can look at profiles across multiple assays, such as a heat map color coded according to AC_{50} values, which spawned a graphical representation of actives, providing another visual method of investigating correlations. He showed another visual representation, which illustrated the NTP decision tree. Sometimes finding correlations between profiles is of interest, whereas sometimes finding differences can be informative. He showed a pair wise comparison between assays that illustrated that point.

Dr. Shockley briefly discussed the Tox21 toxicity databases. ACToR, the EPA's Aggregated Computational Toxicology Resource, contains 500,000 environmental chemicals. EPA also offers ToxRefDB—the Toxicity Reference Database contains almost 2,000 *in vitro* pesticide registration toxicity studies involving hundreds of chemicals. The NTP supports CEBS with extensive information on thousands of chemicals including NTP historical data. Each database is searchable, and offers specific features of interest.

With regard to future plans for the IWG, Dr. Shockley noted that as Tox21 moves into Phase II, with its 10K library, new algorithms for making activity calls would be needed. The current activity call algorithms will be implemented into CEBS for public use. The group will determine optimal approaches for evaluating differential activity, and will integrate data from different assay types for more extensive toxicological analysis and to better prioritize compounds for targeted testing. Also, the IWG will develop methods for predicting *in vivo* human responses after integrating data from toxicological databases.

In summary, Dr. Shockley said that the ability of a substance to induce a toxicological response is better understood by examining responses at multiple concentrations. He mentioned that making activity calls for numerically large compound libraries is complex and depends on the underlying data structure and the focus of the study. Prioritization requires integration of data across different assay formats involving different approaches for identifying active compounds. New informatics tools will be required to deal with the complexity of using *in vitro* data to predict *in vivo* toxicity. He said all Tox21 tools and data would be placed in publicly accessible, integrated databases.

B. BSC Questions

Dr. Nicholas Jewell asked whether the method of removing outliers substantially improved the results of the algorithm. Dr. Shockley answered that it did. Dr. Jewell asked whether the system of classifying compounds as active, inactive, or inconclusive is strict, or whether there

might be a place for "fuzzy" classification, to allow for some compounds being more robustly active than others that might be "borderline." Dr. Shockley said the current system is based upon p-values to maintain a specific false discovery rate, and that particular threshold could be lowered to be more conservative. He added that as the project moves forward, only compounds with a high degree of confidence in the activity call would be used for prediction modeling. Dr. Jewell asked about the clustering system Dr. Shockley had shown previously; whether it was done qualitatively or through an unsupervised clustering algorithm. Dr. Shockley replied that hierarchical clustering had been used.

Dr. Sherley agreed that the approach of measuring 15 different concentrations the chemicals is "really valuable," but wondered about that value compared to its effect on throughput. He asked about the gain in sensitivity and specificity from using that method. He also asked about the range of assays, and how many actives might be captured simply measuring cytotoxicity. Dr. Shockley said the value of multiple concentrations is a point of interest, but is difficult to assess given the data points involved, and although it is being worked on, there is no number to attach at present. Regarding Dr. Sherley's second question, he said that it would differ with each assay, given the range of toxic responses involved.

C. Presentation: The NCGC BioPlanet

As part of the IWG's presentation, Dr. Ruili Huang, NCGC, briefed the BSC on a new tool called the NCGC BioPlanet, a visualization method incorporating all of the known pathways and annotated with a great deal of related information. The pathway annotations were taken only from manually curated, public sources. With pathways integrated from a number of different data sources, it annotates pathways by source, species, biological function or process, disease or toxicity relevance, and assay availability. It provides easy visualization, browsing and analysis of pathways, and facilitates pathway assay selection and prioritization in Tox21. The BioPlanet will be made publicly available in early 2011.

Switching to live interaction with the software itself, Dr. Huang demonstrated many of the BioPlanet's features. She showed, for example, that each "star" on the planet represented a different pathway. Different filters can be applied to access various levels of information, including assays, disease pathways, toxicity pathways, and PubChem information on pathways with available assays, plus pathways with ToxCast™, Tox21, and NCGC assays. By clicking on a pathway and choosing various features, annotated information from the literature can be displayed, including the genes associated with a pathway. Search by gene can also be conducted, to see all pathways containing a particular gene, or multiple genes, such as autism, diabetes, or obesity genes. Search by disease is also possible.

Dr. Huang said the BioPlanet is a useful tool for pathway prioritization in Tox21, with its ability to visualize toxicity pathways, disease pathways, and assay availability from PubChem and commercial assays. It will also aid identification and development of new assays for pathways with no current assay coverage. Future developments for the project include the addition of links to compound activity data, the incorporation of other data forms such as sequence data or

gene/protein expression data, the inclusion of pathways from other species, and the organization of assays according to pathways, diseases, and toxicity endpoints.

XIV. Tox21 Working Groups: Targeted Testing Working Group

A. Presentation

Dr. Michael DeVito, NIEHS/NTP, presented information about the Targeted Testing Working Group (TTWG). The goal of the TTWG is to build a bridge between HTS results and risk management decisions by evaluating the qualitative and quantitative relationships between *in vitro* HTS assays and predictive models to *in vivo* biological activity and toxicity. He said the group is quite collaborative with the other working groups, but is currently somewhat behind the others, needing to await results in order to proceed. Currently, the group is evaluating some of the predictive models generated by the ToxCast™ data, as they are the models now available.

Dr. DeVito reminded the BSC that there were 320 ToxCast™ Phase I chemicals, with 309 unique structures, and that they were predominantly pesticides. He also briefly reviewed the ToxCast™ array of assays, both biochemical and cellular, which resulted in more than 500 endpoints.

The first model being evaluated by the TTWG is a liver targeted testing study involving a statistical model developed by NCCT that predicts rodent liver proliferative lesions and rat liver tumors, based on ToxCast™ Phase I screening data. It was developed using multivariate analysis for a subset of 21 ToxCast™ chemicals with positive rat liver tumor findings. The model was applied to the ToxCast™ data set to evaluate how well it would predict the results from the other chemicals.

Dr. DeVito related the assays associated with rat liver tumors in order of statistical significance. The first, a PPARγ transactivation assay from Attagene, is a prerequisite—to be a rat liver carcinogen, this assay must be activated. Also, the chemical must be active in one or more of the following assays: PPARα, anti-AR, induction of HMGCoASII, oxidative stress, and induction of CCL2.

Sixty-nine of the 309 ToxCast™ chemicals were predicted to be non-genotoxic rat liver carcinogens. Six of the 69 had no data, and five were considered FPs. Eighteen of the 21 rat liver tumorigens were identified with this model—three that were not identified had no activity in any of the assays, but were assumed to be false negatives due to stability or solubility issues, not that they were biologically inert. Dr. DeVito showed a table depicting the results for the 69 chemicals, which illustrated the fact that the predictive model was relatively successful.

Dr. DeVito described the inherent uncertainties associated with the model: (1) the use of human-based cell assays to predict rodent tumors, (2) the question of whether the *in vitro* assays actually represent *in vivo* endpoints, (3) the lack of metabolism in many of the *in vitro* assays, and (4) the question of what is a hit *in vitro* and how does it relate to *in vivo* responses. There is also some uncertainty in the model, he added. First, the markers are somewhat unexpected, in that most experts would expect the prerequisite marker to be PPARα, not PPARγ, in that much data suggest that PPARα drives tumors. Also, AR antagonism is

apparently involved, but in humans, AR antagonists are chemotherapeutic agents for liver cancer. Chemokine (C-C motif) ligand 2 (*CCL2*), a chemokine involved in chemotaxis and angiogenesis, appears to be a likely candidate. Finally, there is oxidative stress (OS) as represented by double-stranded DNA breaks, which is anomalous in that the compounds are classified as non-genotoxic. Some of the chemicals are known to activate PPARα *in vivo* and induce liver tumors, but for most chemicals there is limited or no information on their *in vivo* effects consistent with the observed assays.

The goals of the liver targeted testing project are (1) to test for the *in vivo* presence of activity seen *in vitro*, looking for sensitivity, specificity, and dose-response; (2) to confirm that previously untested compounds show predicted *in vivo* activity; and (3) to see if a Reverse Toxicokinetics (RTK) approach gives a reasonable estimate of dose for *in vitro* to *in vivo* extrapolation

The project has two hypotheses: (1) i*n vitro* activation of PPARγ, along with one or more of the CCL2, AR, OS, or PPARα pathways, is highly predictive of the corresponding activation *in vivo*, at some dose level and (2) only at doses for which at least two of these pathways or processes are activated will liver tumors be induced in the 2-year rat study.

To test these hypotheses, said Dr. DeVito, a tiered approach will be employed. Tier 1 is pilot studies evaluating numerous chemicals at one dose and time point, investigating whether at the highest dose tested in a bioassay, *in vivo* signatures consistent with *in vitro* results are seen. Tier 2, depending on the results of Tier 1, consists of *in vivo* time course and dose response studies, including RTK. Tier 3, depending on the results of Tier 2, evaluates chemicals without 2-year bioassays.

Providing more details on the Tier 1 pilot-screening project, he noted that it would be an iterative process, starting with 12 chemicals and perhaps going as high as 40. The test will be a single, daily exposure by oral gavage to the highest dose used in the 2-year bioassay, for 4 days, with sacrifice 4 hours after the last dose. Tests will measure a variety of toxicity markers. Dr. DeVito also detailed the specific assays that will be used to screen for PPARα, PPARγ, and CCL2 activation, oxidative stress, and AR antagonism.

Initial Tier 1 chemicals will be selected according to specific criteria: they must be included in ToxCast™ 309, have 2-year bioassay data, have been predicted by the model to be positive or negative, and have been tested in male Sprague-Dawley rats, which showed the largest number of chemicals positive for liver tumors in a single strain or gender.

The list of initial chemicals includes several predicted to be positive by the model and shown as positive in the ToxRefDB, one true negative, and two chemicals that were negative in rats but are predicted to be positive by the model. Dr. DeVito pointed out that the listed chemicals cause rat liver tumors at most one out of five times, thus it is not unexpected that the model would yield a number of FPs.

In terms of significance and expected outcomes, *in vitro* assay signatures will be compared with *in vivo* data, potential pathways and targets for future development will be identified; the group

will develop tools to extrapolate *in vitro* concentrations to *in vivo* doses and exposures, and will provide feedback to the IGW and APWG.

Dr. DeVito described the RTK approach in more detail. It uses *in vitro* and computational methods to predict the exposure that results in blood concentrations equivalent to the media concentrations, assuming that media is equivalent to blood. It takes advantage of advances in predictive pharmacokinetics and applies them to toxicological questions.

One remaining question about RTK is what the media concentration represents, qualitatively (e.g., blood, plasma, serum, or tissue) or quantitatively (a direct 1:1 relationship or proportional). Dr. DeVito said the present assumption is that if media equals blood, then the ratio between media and cells equals the ratio between blood and tissue.

Dr. DeVito mentioned that the TTWG is also considering projects involving EDs including one with ToxPi prioritization.

In summary, Dr. DeVito said the TTWG involves ongoing efforts to evaluate relationships between HTS results, predictive models, and *in vivo* effects, both qualitatively and quantitatively. He said initial efforts include evaluation of a predictive model for non-genotoxic liver carcinogens, and that efforts are beginning to develop an endocrine disruptor project. He said the group's efforts should provide insight into the uncertainties in extrapolation of HTS data to *in vivo* biological and toxicological responses.

B. BSC Questions

Dr. Miller asked about the graphic depicting the initial chemicals and why PPARα was not included in the ToxCast™ results list. Dr. DeVito replied that on that list, PPARα had been represented as HMGCS2, the gene up-regulated by PPARα.

Dr. Teeguarden asked which working group was ultimately responsible for the call regarding when rank ordering *in vitro* predicts rank ordering *in vivo*. Dr. DeVito responded that he believed it would be a combination of the TTWG and the IWG working together. He said it would also involve the APWG, but that targeted testing would lead the work. Dr. Teeguarden inferred that there is no one person responsible, i.e., no "czar," to which Dr. DeVito agreed.

Dr. Minor asked about the use of extrapolated exposure levels in the liver toxicity studies. Dr. DeVito said the studies hadn't actually been started yet and the exposures were estimated based on dietary exposures in the bioassay. Dr. Minor was confused about the fact that results were being shown, if the studies hadn't been done. Dr. DeVito clarified that the results shown were from bioassays that had been done by the manufacturers over the past several decades. The ToxCast™ results that were shown were all *in vitro* and were shown to illustrate the question of correlation with the *in vivo* tumor results. Dr. Minor asked about plans to take blood samples to determine compound concentrations when the studies are conducted. Dr. DeVito replied that it would be done in Phase II of the testing, since the early testing would only involve one dose and one time point.

The meeting participants attended the BSB poster session in the lobby.

C. Tox21 Working Groups: BSC Discussion

Mr. Janzen, first lead reviewer, said the CSWG had very adequately addressed the goals of Tox21, carrying the initial phases of Tox21 forward toward building a well-characterized set of compounds, and to build on the known toxicity profiles from Phase I. With the goal of building the 10K compound library, they will be able to build a database that is quite comprehensive. He was happy with the thoughtful layout of the plates and applauded the group for gaining access to the 100 failed pharmaceutical compounds. He pointed out that the NCGC library was designed to focus on druggable compounds, and there may be fewer toxicity hits in it. He was unclear about the logistics of how the sets were being created, and noted that it was not important that a single vendor prepare all of the compounds as long as a unified process was employed. He was pleased with the QC being applied.

Regarding the APWG, Mr. Janzen felt that the group's goals and responsibilities were clearly articulated and align well with the Tox21 mission, with a strong group and a strong group of advisors. He concurred with the plan to use the qHTS approach to take the project forward. Without going into individual assays, he mentioned that it would be important for many of them to move into the 384-well format, and that the group should carefully balance cost and time versus productivity in the choice of which assays to miniaturize to the 1536-well format. He said that he does not approve of the use of signal-to-background as a test of assay quality and urged the group not to use it as a pass/fail measure. He favored the SBIR programs being supported by the group.

Mr. Janzen approved the work of the IWG, praising their efforts in working across assays, vendors, and platforms. He expressed concern about the method of dropping outliers in the curve class calculations; there should be a limit on how many outliers can be dropped. He encouraged the group to link as strongly as possible their work on human prediction into the existing FDA QSAR models.

He felt that the TTWG presentation was put together quite well. He felt that the media/cells=blood/tissues concept was probably not a good starting point, and that dose binding studies might be a better method for initial calculation.

Generally, he felt that the working groups had made a very strong showing, and that they had moved forward "incredibly" since the initial meeting in 2005.

Ms. Ruthann Rudel, second lead reviewer, echoed Mr. Janzen's admiration, calling the presentations "very exciting and a great start to the program." She felt this is a good time for the program to express explicit, achievable, reasonable goals and expectations. She said whole animal testing would always be needed, especially for risk assessments for important chemicals, and so is reluctant to see Tox21 as a replacement, particularly because there are some significant challenges that may never be overcome, such as cell type, tissue type, the timing of assessment, serum context, solubility, metabolism, and volatility. With so many variables, she said, the answer to the question "Is it safe?" will always be, "It depends on how it was tested." To be realistic, the limitations should be kept in mind, and not to say there won't be progress in many of those areas.

She called for demonstration of some specific applications of the data in the near term. For example, providing data where we now have none, given that whole animal testing is expensive and for many chemicals there are no data. She wondered whether Tox21 could be used in product formulation decisions, where there are currently inadequate data. She suggested the Tox21 HTS data might be useful to allow a choice of just one targeted animal study, skipping others. The data could be used to design occupational epidemiological studies with a sensitive early effect marker. Tox21 could be useful to assess mixtures, which have been such a difficult issue for toxicology to confront, including the assessment of herbal preparations and other natural products. She wondered how the HTS data could be demonstrated to be superior to current *in silico* models, or be used to improve those models. She acknowledged that the application projects she was suggesting would require expanding the scope of the working groups, but that they would be a platform for public communications about the project.

She liked the disease-specific pathways approach, but wondered about the implications for the choice of cell types used in the assessments; there will eventually be a need to expand the number of cell types, as opposed to the desire to limit them at this point. She was unclear about the current rationale for including different tissues or cell types.

Dr. Sherley, third lead reviewer, said Tox21 is "absolutely something we should be doing." The scope is outstanding, but he cautioned that it would be challenging to triage the appropriate use of resources. He felt that this is a moment at which the mindset of the parties involved needs to change, with the coming major expansion. For example, he wondered, what is an acceptable FN rate in the larger scope?

In terms of approach, he said testing assumptions would be important as the move is made from 1,000 to 10,000 compounds. He noted that the choice of assays for the start-up of the program has been based upon available, accessible assays, but that may not be *a priori* the best set to be using, and that additional assays may need to be developed.

Dr. Teeguarden, fourth lead reviewer, said he was "deeply critical across the board" compared to the other reviewers. He said the program is "absolutely extraordinary," and he is certain there will be incredible breakthroughs as a result. He felt that the program has everything needed to develop predictive toxicology in terms of breadth and depth, collaboration, and leadership, and that there is every reason to believe it will be successful.

Dr. Teeguarden began his general programmatic comments by again asking where the final responsibility lies for reaching the ultimate goal of Tox21—hazard rankings *in vitro* that predict hazard rankings *in vivo*. He said apparently that responsibility is distributed, and that will not work. He perceives too much reliance on faith that massive data sets are going to lead to the right answer, and not enough proof of concept, akin to the overselling of other technologies such as genomics, which took a long time to achieve high value. He expressed concern about timing, with a lack of timelines and milestones articulated, leaving the program vulnerable to tough questions about productivity. The program should not be left to suffer in fear about not meeting its ultimate objective, because there are many scientific advances to be made before it can be achieved. He called for each of the working groups to have achievable, step-wise

milestones in place, leading to the ultimate achievement. He also wondered who is responsible for the integration of the working groups.

Regarding the APWG, Dr. Teeguarden felt that it had a premature emphasis on the development of high throughput methods, with assays being chosen largely on the basis of whether they would work in the high throughput format, rather than their predictive value. He felt that the assays should be chosen from the "bottom up, the biology up," and should be selected for sensitivity, specificity, and rank order potency. He urged consideration of middle ground; medium throughput assays that may be useful, and could help drive innovation in the development of high throughput assays.

He felt that the IWG needs stated objectives and milestones, and needs to focus more on validating rank ordering. He wanted to make sure that the CSWG and the IWG are actually working closely with the TTWG. He wanted to see stated decision points about how and when the decision is made about correlations between *in vitro* and *in vivo* outcomes. He recommended adding some common nutrients such as glucose or vitamins into the mix of chemicals. He also mentioned that relative to some of the systems in the program, in 20 years a rat study might appear to be simple and effective.

Dr. Minor, fifth lead reviewer, noted that it was her first BSC meeting and said she was impressed with the Tox21 program. Regarding the CSWG, she said it is important to develop a well-characterized compound library. Without one, it is impossible to make any conclusions about anything else being done. She felt that the 10,000-11,000 compounds are probably adequate to start, but that more may be needed to allow focus on key toxicity areas that may not presently be included. She felt that DMSO is "unfortunately" the most practical solvent to use at the start. She approved of the QC procedures described, but was concerned about the 90% acceptable purity figure, with the issue of potential impurities. She approved of developing a water-soluble compound library, and suggested that compounds soluble in both water and DMSO should be included. The inclusion of the failed pharmaceutical compounds is of great value, she said. She was concerned, however, about the lack of procedures to confirm the identity and concentration of compounds in the plate wells, as those details should be documented. She also advocated inclusion of defined mixtures of compounds, as opposed to undefined mixtures such as natural products.

Regarding the APWG, Dr. Minor liked the HEPG2 cells approach, but was concerned about the choice of culture medium for those assays, due to issues related to mitochondrial activation. For the NR assays, she noted how complicated the receptors are, and advocated the use of full-length versus truncated receptors, as they could provide pieces of the puzzle otherwise left missing. She approved of the ToxCast™ Phase II screens, particularly the 3D tissue models, and asked what was being done to ensure cell quality before and during the screen. She questioned whether the ingredients of the culture mediums were known, as many vendors do not provide that information, and inquired whether assays for assessing mitochondrial DNA damage have been considered.

Regarding the IWG, she felt it had taken a good approach to a daunting challenge. She approved of their practice of not trying to just develop their own systems but also look at what is publicly available. She recommended development of a central repository for publicly generated data linked to each compound or chemotype. She liked the BioPlanet, and wondered if there was a plan to link to the references that had generated the data.

Regarding the TTWG, Dr. Minor was concerned about the methodology, citing her own background in doing cell-based assays and trying to correlate the results with *in vivo* results. She recommended seeing which compounds were active in the cell culture or biochemical assays, and then doing a pharmacokinetics analysis in animals to measure blood levels, determining the highest blood level prior to the targeted testing. She said the current method is like walking before crawling, and they should crawl first. She also wondered whether the proposed compounds or proteins had been tested in the test species in terms of rank order potency. Finally, she expressed concern about looking at the breadth of the data for chosen compounds across the different assays, in that some observed toxicities might be due to off-target events.

Dr. Tice responded by saying that many good points had been raised. He said Tox21 had in fact paid special attention to inclusion of negative compounds in the libraries. Responding to Ms. Rudel's comment about the articulation of goals, he reminded the BSC about the matrix he had presented, which included goals that drove the activities of each of the working groups. He said in the next day's re-visitation of the chart, he would provide information on accomplishments and ongoing projects, to be able to put the goals into perspective. He also pointed to his presentation as an articulation of the value of HTS data to NTP, and by extrapolation to the other organizations. He cited the example of green chemistry, and how to decide which chemicals are greener than others. Regarding mixtures, he said Tox21 is clearly trying to work with them, recognizing the difference between defined and ill-defined mixtures, such as herbals. The ability to run large numbers of compounds in qHTS makes studies on mixtures possible.

Regarding the discussion of pathways and tissues, he said the choices had been made based on an understanding that some chemicals act differently in different tissues, and were supported by the EPA's ToxCast™ testing across 500 assays. He said Tox21 looks at NCGC as the first stage, to identify compounds that interact with a specific target, to be followed by looking at more specific targets reflective of tissue differences. It is difficult to use primary cells in a 1536-well format, so primary cells or stem cells are considered as the secondary level of assays.

Responding to Dr. Sherley's question about the acceptability of FNs, Dr. Tice mentioned that they are a concern, but that the Tox21 approach is based on using batteries of assays, some of which would be orthogonal. This approach might help to reduce the overall frequency of FNs as well as FPs.

Regarding Tox21 decision-making responsibility, he said "the buck stops" with the points of contact from the four agencies, who interact biweekly, interact monthly with the working group co-chairs, and meet quarterly with all staff involved in Tox21 activities to decide how or whether

to proceed with various aspects of the program. There is a group at NIH that implements the Government Performance and Response Act (GPRA) and oversees performance of large initiatives. Tox21 has articulated annual milestones in its reports, including how many compounds have been tested in how many assays, with the ultimate goal of establishing a prioritization scheme that works.

He noted that there had been much focus during the working group presentations on the qHTS studies conducted at the NCGC because that is the component NTP is focusing on, while EPA focuses on ToxCast™, and stressed that the goal is to integrate, not duplicate programs. He added that although a newly nominated compound may not be added to the library, it could still be addressed through targeted testing, for example, or by adding compounds to ToxCast™. Thus, there are strategic testing capabilities, and under discussion is the ability to conduct targeted testing within a contract or perhaps in the NTP laboratories that are being established.

Responding to the question about the 90% purity cut-off, he noted that it is not a cut-off but a value, and that impurities are always going to be a concern; that is why secondary follow-ups are always conducted. He also addressed the question of how much compound is in a well, and said that is an issue they have been wrestling with for a long time. He said the free concentration of the compound in a well is what is actually important, and that ways of addressing that are being explored.

Regarding HepG2 cells, he said it is clearly understood that growth on glucose affects mitochondrial activity, and there is already a plan to conduct a follow-up study with HepG2 cells maintained on galactose; he directed the BSC to one of the posters that provided information on mitochondrial toxicants detected at the NCGC. He stated also that NTP has purchased a Seahorse instrument specifically to evaluate mitochondrial toxicity *in vitro* and *in vivo*.

He said there was no link to sources in BioPlanet. Dr. Austin added that it would be a hugely complicated undertaking to add that capability.

Dr. Tice noted that the addition of DrugMatrix would be a major step forward for linking pathway data to disease.

Dr. Novak closed the session, stating that the response by the BSC had been “highly positive,” with an exceptional group of resources brought to bear on the problem, and having enormous potential for major contributions to predictive toxicology and human health.

December 1, 2010

REVIEW OF THE BSB (continued)

XV. Tox21 Activities: Introduction

Dr. Novak convened the meeting, participants introduced themselves, and Dr. White read the conflict of interest statement. Dr. Bucher noted that the review of the BSB was to continue, focusing on its role as the major NTP participant in the Tox21 initiative. He said this session

would concentrate on several activities that support Tox21, but also have broader applications for the BSB and the NTP.

Introducing the Tox21 Activities portion, Dr. Tice thanked the BSC members for their interesting and useful comments from the prior day's proceedings. He returned to the slide he had shown previously, depicting the program's concepts, and illustrating how the working groups and activities relate.

XVI. Tox21 Activities: NTP Caenorhabditis Elegans Screening Facility (WormTox)

A. Presentation

Dr. Jonathan Freedman, NIEHS and Head of the WormTox Group, briefed the BSC on the WormTox research activities. He began with some background about the model organism, *Caenorhabditis elegans (C. elegans).* It is a non-parasitic nematode, approximately 1 mm in length, with a 10-day lifespan. It has highly differentiated digestive, reproductive, muscular, and nervous systems, and its cell lineage is known for the entire development cycle. Transgenic nematodes are easily generated and it is amenable to classic and molecular genetic analysis, with its small genome (100 Mb), which has been completely sequenced. With current sequencing technology, 40-fold coverage of the entire genome can now be accomplished in ten days. Dr. Freedman illustrated the *C. elegans* development cycle.

One of the strengths of the model is the amount of conservation between it and mammals, including basic metabolic proteins, stress response, cell cycle control, several signal transduction pathways, and many neurotransmitters and neuroreceptors. It is highly useful as a model for human diseases such as cancer and several neurodegenerative diseases.

The WormTox project was originally initiated in 2004, with five identified tasks: (1) develop methods to measure the toxicity of developmental and neurological toxicants, (2) expose *C. elegans* to at least 200 known or suspected developmental and/or neurological toxicants, (3) create and/or obtain GFP-based, stress-responsive *C. elegans* for improving sensitivity and specificity of toxicity screens, (4) use *C. elegans* microarray analysis and test a subset of chemicals from Task 3, and (5) adapt methods for high throughput analysis to assess the toxicological responses in *C. elegans* in which each gene has been inactivated using RNA interference. All of those tasks are basically completed except Task 4, which was eliminated as being too expensive and labor-intensive.

There are two groups of assays used. Medium-throughput assays involve five chemicals per week, assaying reproduction, feeding, growth and movement. High throughput assays measuring reproduction and growth are performed for about 100 chemicals per week. Reproduction, feeding, and growth assays use the Complex Object Parametric Analyzer and Sorter (COPAS) Biosort, a fluorescence-activated worm sorter that can dispense *C. elegans* in exact numbers at specific developmental stages. The instrument works in a 96-well format. Dr. Freedman showed two examples of data plots from assays run on the instrument. He outlined a list of the chemicals that have been examined in WormTox, many of which are metals. The

Tox21 libraries have also been run through WormTox—the ToxCast™ 320 library, and the NTP 1408 library.

Dr. Freedman described the *C. elegans* toxicity assays in more detail, outlining the protocols for the growth, reproduction, and feeding assays. He described HTS using the growth assay, for analysis of the ToxCast™ Phase I chemical library. In the WormTox screen of the ToxCast™ 320, 50 L1-stage nematodes were loaded in each well of a 96-well plate using COPAS. They were exposed to one of seven concentrations of toxicants and incubated for 48 hours at 20°C. Size and population distribution were then analyzed using the COPAS. The chemicals' toxicities were ranked according to dose-response, using a variety of metrics to define toxicity in the worms. Ultimately, compound activity scores were chosen as the best metric. In the first screen, it was determined whether the chemicals were active at 200 µM, the maximum concentration tested. A chemical was defined as active if there was evidence of severely retarded growth. If there was some growth retardation, but some overlap with control, the compound was deemed inconclusive. If growth did not differ from control, the chemical was considered inactive. He showed data illustrating the results of the assay, depicting decrease in size over time. Chemical activity is also determined by calculating trends using dose response for seven concentrations. With this method, compound activity scores were determined, assigning an activity score based on summing activity results, using a point system based on decreases in size per concentration and slope. Compounds with an activity score of 2-9 were deemed active, a 1 considered inconclusive, with a 0 considered inactive.

Dr. Freedman showed a pie chart illustrating the activity scores of the ToxCast™ compounds at the highest concentration. There were 150 actives (48%), 113 inactives (37%), and 46 inconclusives (15%). Of the actives, 35 compounds had activity scores between 6 and 9, indicating high toxicity. Of those 35, 23 were insecticides, seven were fungicides, three were herbicides, and one was a microbicide.

Another step following the ToxCast™ 320 screening has been to compare results with other species, particularly to data collected in a zebrafish developing embryo model system by an EPA collaborator. Dr. Freedman showed a comparison of the different experimental conditions used with the two species. Although in-depth analysis of the data has not yet been conducted, initial study indicates (after removing the inconclusive compounds from consideration) 80 inactive compounds in both species, 128 actives in both species, and 22 and 32 active solely in *C. elegans* and zebrafish, respectively. Sensitivity, specificity, positive predictive value, negative predictive value, and concordance were all at high levels. Dr. Freedman showed a heat map illustrating the concordance within chemical classes.

He said the next major project for WormTox involves the screening of compounds in stress-responsive transgenic *C. elegans*; these transgenic worms are being produced through a SBIR contract. The contractor will generate three lines per gene target, with each line containing one of 34 different genes selected for a variety of stress response characteristics. The data will be analyzed by a profiler attachment to COPAS, which will measure fluorescence, corresponding to outcomes.

Future plans for WormTox include comparison of *C. elegans* ToxCast™ data to other *in vivo* and *in vitro* ToxCast™ Phase I datasets. The data will be made publicly available through CEBS. Results will be used to determine the design of future screens of larger chemical libraries, such as ToxCast™ Phase II (700 chemicals) and portions of the Tox21 10K library.

B. BSC Questions

Dr. Elaine Faustman said she noticed that the metals are positive in *C. elegans* and asked about the kinetics and uptake. She mentioned the similarities in pathways between *C. elegans* and humans, but noticed that there were no hormone receptors listed, despite the strength of the model in reproductive endpoints. Additionally, she inquired about metabolism in the assays, not seeing oxone versus chlorpyrifos (CP) listed.

Dr. Freedman said they had tested oxone versus CP directly. Dr. Faustman asked if CP was positive without the oxone. Dr. Freedman said it was, and that *C. elegans* will metabolize many of the chemicals, although the comparison between metabolism in *C. elegans* and higher organisms is not known. Dr. Faustman said that it is a strength, and should be included in Dr. Freedman's slides. Dr. Freedman said in terms of the target genes they had selected, they were still developing the transgenic worms, and would welcome the BSC's suggestions for target genes. Regarding the kinetics of the metals in the assay, he said metal uptake had not been analyzed, but it appears that uptake of metals is rapid. He said that question was one of the reasons they had developed the feeding assay.

Dr. Teeguarden asked how WormTox interacts with the Tox21 working groups, particularly the TTWG. Dr. Freedman said they already work with the Assays group, and need to talk with Dr. DeVito and the TTWG about how to incorporate the *C. elegans* data, and offer their services for supplemental testing.

Dr. Zelikoff asked how the system might handle inhaled material or insoluble materials or particular matter, or if it is limited to soluble materials. Dr. Freedman said that as long as the material is smaller than the animal's mouth, it would be ingested. He said the animal lives in an aqueous environment, so unless the volatile material can be dissolved into that aqueous layer, it would not be absorbed. Dr. Zelikoff speculated that as a result, *C. elegans* would not be a good model for inhaled chemicals, particularly since the lung often mediates the toxicity of inhaled materials. Dr. Freedman agreed.

Following up on Dr. Teeguarden's question, Dr. Bucher said WormTox is an intermediate stage. He asked Dr. Freedman whether it was more suited for confirming strong responses in the HTS assays, or weak responses, and whether he had had an opportunity to sort through the data to begin to address those questions. Dr. Freedman said they have begun that assessment, and that in a few months he would be more able to answer the question.

Ms. Rudel asked about the comparison of data between *C. elegans* and zebrafish, in that the growth assay was used for the zebrafish, but growth and development assays were used for the worms. Dr. Freedman said the growth assay was all that was available for the zebrafish, and that determining how to compare the endpoints was part of the challenge. Noting that the data

shown have been associated with high doses, Ms. Rudel wondered about what had been seen at lower doses. Dr. Freedman clarified that the activity scores were actually reflective of the dose response. Ms. Rudel asked how an increase in exposure might affect the scatter in the growth curve measurement. Dr. Freedman said the response depends on the chemical. She said that might be an interesting parameter to include in the data analysis. Dr. Freedman replied that they actually have done so, and are constantly looking for new ways to analyze the data.

Dr. Teeguarden noted that one of the limitations of HTS is the ability to use a very limited number of cell types, and that WormTox is able to use an intact organism with multiple, well-characterized cell types, and now transgenic models as well. He asked whether in the transgenic worms there is a way to tie the results back to multiple cell types, particularly over time, including mechanism of action, triggered genes, etc. Dr. Freedman said yes, that is possible, but the question is how to do it. Dr. Teeguarden suggested 3D or 4D mapping of all of the worm cell types. Dr. Freedman said the issue with that proposal is throughput; time and resources do not currently allow for that level of detail on a large scale. He said the group is aware of the power of that type of analysis, particularly looking at individual animals under a microscope to see which cell types are expressed, time course, and other data, and then linking that information back to data from the profiler. However, the difficulty of performing that analysis rapidly limits its utility.

XVII. Tox21 Activities: Probing Mechanisms of Inter-individual Susceptibility to Toxicants with Population-based Experimental Approaches

A. Presentation

Dr. Ivan Rusyn, University of North Carolina at Chapel Hill (UNC), a NIEHS/NTP collaborator, briefed the BSC on his group's work connected to Tox21. He said he and his collaborators are trying to work in one section of the new paradigm for toxicology by looking at mode of action of various chemicals, dose-response analysis both in individual genes and pathways, and also starting to think about how to incorporate population into *in vitro* testing.

Since the initial NTP-NCGC qHTS screening study published in 2008 outlining the use of HTS to screen chemicals, Dr. Rusyn's group in collaboration with Dr. Alex Tropsha's group at UNC has used that study's publicly available data in two subsequent publications to elaborate on how it could be made useful for toxicology and decision-making. Capitalizing on this progress, they considered how population-based studies could be incorporated. They are "a reality across all model systems" at this point, including epidemiological studies, *in vivo* animal studies, and *in vitro* studies. In humans, it is now common practice to collect hundreds of single nucleotide polymorphisms (SNPs) from thousands of individuals and do genome-wide association study (GWAS) analysis. In animal studies, population models of rodents have been built (largely by NTP) to capture the genetic diversity of populations, with complete sequences of dozens to hundreds of animal strains. In *in vitro* studies, there are now hundreds to thousands of cell lines available that have been deep sequenced or genotyped.

Looking to the possible future relationship of *in vitro* testing to risk assessment, and how it might change the paradigm from the current *in vivo* model, Dr. Rusyn said an *in vitro* human cell-based model fills critical gaps regarding hazard identification, mode of action, dose-response, and variability analyses. In this paradigm, a genetically diverse population can be defined by genotyping and put into an *in vitro* model system, leading to its use in toxicity testing, as opposed to drug efficacy testing as has been common up to this point.

The first experiment by Dr. Rusyn's group focused on hazard identification analysis. It consisted of one of the first collection of lymphoblasts representative of a sample population collected and haplotyped by the International HapMap project. These comprise 87 lymphoblastoid cell lines (29 families, parent-child trios), a large, renewable resource with publicly available information. These cells are easy to manipulate and control, are representative of genomic DNA from a diverse human population, and have been densely genotyped (>$5x10^6$ SNPs), allowing for association mapping of the phenotypic differences between subjects. The experiment was low-throughput, in 96-well plates. Fourteen chemicals were tested in three concentrations on 85 cell lines. Two assays were used, one that evaluated cytotoxicity, based on the measuring levels of adenosine triphosphate (ATP), and one that measured caspase-3/7 activation, an endpoint indicative of apoptosis. The study created 14,280 data points. Dr. Rusyn said hazard identification comes from being able to screen cytotoxicity in different cell lines from different individuals, but that inter-individual differences in responses are also important and informative. He showed data from screens with three chemicals that depicted differences in the extent of cytotoxicity among individuals. Particularly with perfluorooctanoic acid (PFOA), the data showed considerable variation in individual responses. Such data not only capture the ability of chemical to be cytotoxic, but also capture the degree of inter-individual variability in response.

Dr. Rusyn acknowledged that the question of metabolism is present with these cell lines, although the lymphoblastoid cells are not completely devoid of metabolism. He considered a lymphoblast likely to be as metabolically proficient as a primary hepatocyte that had been in culture for 3-5 days. Both genetic and non-genetic factors influence responses in the lymphoblastoid cell lines. One of the challenges is significant day-to-day variability in responses, which can be controlled for. Another challenge is that it has been suggested that responses to different chemicals may correlate to each other because they correlate to growth rate, ATP metabolism, and other forms of transformation the cells have undergone. There are a number of statistical or experimental methods of controlling for these types of variability. Of the 14 chemicals tested, only two showed significant inter-individual variability in response. This method would be one way to prioritize chemicals for further testing based on the extent of inter-individual variability, Dr. Rusyn pointed out.

Mode of action similarities across a population can also be investigated with this method, to see mechanistically how the mode of cell death and the actual response correlate on the population level. He showed data depicting the concept of how the mode of action information can show which responses are population-level and which are inter-individual.

The data collected in these experiments may not hold sufficient statistical power for a meaningful GWAS, as illustrated by the data he showed for correlations in responses between parents and children, which was poor, showing little heritability. He articulated the advantages in this approach, which include an understanding of inter-individual variability. It allows a genetically defined, genetically diverse human *in vitro* model system to screen for adverse effects of environmental chemicals. Some, but not all chemicals elicit inter-individual variation in responses, so chemicals that vary in their effects across populations may need to be prioritized for further testing. The approach captures a population-wide measure of uncertainty and provides information that may be crucial for future risk assessment approaches that will rely heavily on *in vitro* data. It allows the exploration of potential differences and/or similarities in modes of action between chemicals and allows assessment of population-wide versus individual effects. Despite the inadequate population to conduct GWAS, it is still possible to make a contribution to the link between genetics and adverse phenotypes. By combining toxicity data with publicly available genetic information, it is possible to probe and select candidate susceptibility genes and pathways/networks. Regarding this point, Dr. Rusyn showed data depicting how suggestive genetic associations can be explored in this type of experiment, although he acknowledged that the population is insufficient for the associations to be considered highly reliable.

In another set of experiments, with NCGC, NTP and UNC collaborating, 240 compounds from the NTP 1408 library have been screened in 81 cell lines at 12 concentrations with the ATP-cytotoxicity and caspase-3/7 assays, yielding 1.5 million data points on the population of cells that have been already genotyped for >2-5x10^6 SNPs. He presented data that depicted excellent reproducibility in the responses, showing that the methods employed were dealing effectively with the technical challenges involved. Other data he showed depicted concentration-response and high throughput analyses, with good distribution between actives, inactives, and inconclusives. Further plots showed that some of the chemicals had high degrees of inter-individual variability in response, allowing for follow-up genetic analysis of those compounds. He showed data for these experiments depicting the mode of action similarities across the population, allowing the identification of compounds with similar responses across a population, for both general cytotoxicity and caspase activation.

Dr. Rusyn alluded to the so-called “case of missing heritability,” and said that cell-based studies may help to fill in the gap left in human studies, in which there is disappointingly small correlation between suspected genetic loci and phenotypes. It would be possible for toxicology to leverage the resources of the 1000 Genomes Project in a cell-based system, and that such a sample size, 1000+ cell lines, might be sufficient to detect variants contributing only a few percent of phenotype variation. Other advantages he cited included: (1) toxicity screening data matched by extremely deep ‘omics profiling of the cell lines, (2) the availability of sequencing data will enable the exploration of rare SNP variants, (3) one of the first large datasets to be fully profiled by sequence-based RNA profiling, (4) careful heritability calculations, including population-based approaches that account for “missing” heritability in disease association analyses, and (5) the extensive characterization of these samples will enable dissection of genetic variability. He concluded by noting that the experiments conducted in 90 cell lines could easily and practically be expanded to cover 1000+ cell lines.

B. BSC Questions

Dr. Sherley asked about the missing heritability as it applies to the cell-based experiments, and whether the age or gender of the individuals from whom the cells are extracted is taken into account. He also asked if it was determined whether the cells were all coming from the same tissue source, and whether they are normalized for population doubling. Dr. Rusyn said those are important points to keep in mind, but that the information available from HapMap and 1000 Genomes may be somewhat limited. He said that the cells have all been established from the same tissue source (B-lymphocytes), that the age and sex information on the donors is available, and that isolations and immortalizations are being conducted using standardized protocols. He added that once an experiment is undertaken, all of those factors must be taken into account and can be adjusted for.

Dr. Miller asked about the metabolic competency of the cells, which Dr. Rusyn had characterized as being fairly low, and whether there was a work-around for uptake. Dr. Rusyn replied that if this method works out, and there were particular chemicals to be more deeply explored after HTS, this would be one of the most important experiments to undertake, in trying to understand what the fraction of metabolized compound was in a given screen.

Dr. Faustman asked about the idea of magnifying this approach by expanding it to include study of inter-individual variability in mouse strains, with different organ systems. Dr. Rusyn acknowledged that there are particular limitations in the currently used B-lymphocyte-derived cell lines, and mentioned that his group and some of the collaborators are working with some of the primary cell lines collected from a panel of inbred mouse strains. The basic approach is to try to have a diverse population of cells, but it would be difficult to add the next layer, which would be to have them be fully metabolically competent. He doubted that it would be possible to have that “best of both worlds” in one system in the short term. He said that brings up the question about the usefulness of the data, while knowing that no single model system is perfect. It is important to determine what the data are actually allowing to be predicted. The hope is to take the data to the next level, rather than simply using it to predict population variability. He said understanding the limitations of each model is important.

Dr. Bucher asked whether using a more specific endpoint than cytotoxicity might allow detection of bi-modal populations. Dr. Rusyn said that was possible, but that he was not so pessimistic about the cytotoxicity-ATP assay, because it is highly reproducible and is still very useful. He said in this model system, it would not be able to be used for very deep mode of action analysis, which should come from GWAS and follow-up studies.

Dr. Loomis considered the work to have the potential to bridge the gap between traditional laboratory-based toxicology and human population research, but it raises issues about data analysis and decision-making. He asked how these data might be used to make decisions about which agents are hazardous, especially in situations where there is considerable variation in response across the population. Dr. Rusyn said that when thinking about the overall context of chemical risk assessment and prioritization, one of the critical missing components in the process is population variability, which can be dealt with using defaults, or which can be

measured. The model system will never be perfect, he reiterated, but using population-based *in vitro* model systems could be highly useful in regulatory decision-making regarding which chemicals to scrutinize in depth. It would allow prioritizing based not only on the degree of cytotoxicity, but also on how variable that toxicity is within a population. Thus, he said, it fits a gap in the other approaches that are now being considered, like the EPA's tiered NexGen risk assessment strategy.

XVIII. Tox21 Activities: Mining the NTP Archives for Gene Signatures

A. Presentation

Dr. Alex Merrick, NIEHS/NTP, briefed the BSC on efforts to mine the NTP Archive for gene signatures. The NTP Archive, created in the 1970s, stores more than 2,000 studies, more than 7 million histology slides, more than 4.6 million paraffin blocks, more than 230,000 bags of formalin-fixed tissues, and more than 50,000 frozen specimens. With such a large collection, spanning with work of several decades, the idea arose to pursue the ability to query the samples for gene signatures at the RNA and protein levels.

The NTP decided to derive molecular information from the tissue archive for a number of reasons: (1) the archive is a vast, relatively untapped resource of tissues from toxicological studies; (2) chemically-induced phenotypes (e.g., tumor, organ toxicity) are well-characterized; (3) the data are useful for identifying gene/pathway targets of interest for *in vitro* assays; (4) the archive could be used to evaluate the relevance of *in vitro* prediction models; and (5) genomic changes in chemically-induced tumors in animals could be compared with genomic changes in human tumors, to evaluate species similarities and differences.

The first study to test the idea was a pilot study to determine a reliable method for extracting RNA from formalin-fixed, paraffin-embedded (FFPE) tissues, and then on a gene-by-gene basis, to use quantitative polymerase chain reaction (qPCR) to replicate a gene expression study for which previous microarray data existed. After consultation with colleagues, the NTP determined an effective method for extracting RNA from the paraffin blocks, eliciting material that was felt to be amplifiable. In the first study, they collected material from a recent NTP study that evaluated the hepatocarcinogenic potential of alkenylbenzene flavoring agents, exposing rats to the well-known genotoxicant, aflatoxin B1 (AFB1). The original toxicogenomic study found more than 4,000 significant gene changes after a 90-day exposure to 1 ppm AFB1 in feed, forming a distinct signature, which became a set of gene targets for the extracted RNA study. The changes seen through qPCR were quite similar and comparable to those detected in the original microarray work, suggesting that archival material is, in fact, a queriable and useful source, obviating the need to repeat the original experiment.

Encouraged by that success, the team proceeded to pursue more targeted expression platforms for gene signature validation. The platforms needed to be customizable, with a targeted throughput capacity, using archival NTP FFPE samples, and providing a sensitive signal with small sample requirements. The quantitative nuclease protection assay (qNPA) was chosen. It is similar to qPCR, but multiplexes the amount of information that can be gained. In a 96-well

plate, up to 48 genes per well can be evaluated. Two studies using the qNPA method to validate gene signatures in FFPE archival tissue are now in progress.

Dr. Merrick described Next-Gen sequencing efforts being employed for comparative gene expression in fresh frozen (FF) and FFPE tissues, creating whole transcriptome queries. The objectives of that effort are (1) to explore Next-Gen sequencing platforms as enabling technologies for gene signature discovery in FF and FFPE tissues; (2) to conduct pilot studies using RNA-Seq of FF liver and 3'Seq of FF and FFPE for gene expression profiling for AFB1-exposed liver, and (3) to conduct a pilot study for bisulfate DNASeq to determine methylation sites, to link epigenetic changes to chemical transformation, i.e., to develop the ability to explore epigenetics in the archival samples.

The Next-Gen method RNA-Seq is capable of much higher resolution of the transcriptome than DNA microarrays have offered, due to the use of millions of short reads that are aligned with a reference transcriptome. It offers a >9,000-fold dynamic range compared with 100-fold available with hybridization arrays, as well as other advantages in detection over prior hybridization methods. The BSB first used this approach with its AFB1 samples to see if it would provide more complete coverage of the transcriptome than microarray, and if it would find new transcripts and splice variants specific to AFB1carcinogenesis. Dr. Merrick displayed some of the Next-Gen data that has been generated in the AFB1 study. He also described a new method called 3'Seq, used to sequence the transcriptome in FF and FFPE samples, working around the problem of RNA degradation.

Dr. Merrick mentioned a collaboration with an NTP laboratory that focuses on arsenic and cadmium exposures. The Next-Gen pilot study of aberrant DNA methylation in malignant cells using syngeneic cells transformed in culture is designed to potentially enable epigenetic studies in the NTP archive in the future.

He discussed the recent acquisition of the DrugMatrix database by the NTP. It is a queriable rat toxicogenomic database that includes gene expression profiles, pathology assays, pharmacology assays, and drug literature profiles, along with 637 benchmark drugs and compounds. The diversity of the compound library is one of the strengths of the database. It includes more than 4,000 dose-time-tissue combinations, from approximately 2,000,000 dosed tissue samples. Dr. Merrick presented a chart depicting several aspects of the contents of DrugMatrix, including the fact that the studies had been conducted in multiple tissues, including liver, kidney, heart, rat primary hepatocytes, marrow, spleen, brain and intestine, and muscle, i.e., a very comprehensive range of coverage. Another chart demonstrated the types of queries that can be conducted, which should provide considerable value for the development of predictive toxicology.

In summary, Dr. Merrick said his group believes that NTP archival tissues can be reliably queried for targeted gene expression using qPCR or qNPA, and on a global basis using Next-Gen sequencing to evaluate the relevance of *in vitro* prediction models and to discover new signatures of toxicity. Gene signatures of chemical toxicity will be used to identify critical pathways and improve mechanistic understanding, as well as inform the selection of *in vitro*

assays. He noted that the acquisition of DrugMatrix greatly expands the ability of Tox21 and the scientific community to discover and/or evaluate gene signatures.

B. BSC Questions

Dr. Miller asked for elaboration on initiatives involving epigenetics, particularly as the archival work might link up with the Mouse Methylome Project. Dr. Merrick said he would anticipate that once the mapping of the mouse methylome is complete, it should be much easier to query the archival tissues, as the entire mouse methylome would not need to be explored. He hope there would be a targeted set of genes to query based on the mouse methylome.

Dr. Sherley asked about choices made in the validation of the extraction of the RNA, and plans for moving forward in that area. Dr. Merrick said the choice they had made was to target a specific set of genes altered by AFB1 using FFPE liver from the original NTP microarray study and then validate those same genes from FF liver, considered the gold standard, also from the original study. The very first step was to demonstrate that sufficient high-quality RNA could be extracted from paraffin blocks that then could be amplified and evaluated for consistency with gene changes detected using traditional array-based approaches. The main goal was to establish and validate a set of gene targets with archival tissue. Going forward, the NTP needs to evaluate older archived tissues in terms of quality of extraction and amplification in order to determine how far back in time these samples will remain useful.

Dr. Faustman asked how much tissue was needed from particular structures for the RNA extraction; whether a punch biopsy was used at specified places in the histology, and how big the biopsies need to be. Dr. Merrick described the process of extraction. One tissue per block is used, and the main concern is to have enough material to be able to amplify RNA, with an adequate base pair range to mirror the transcriptome. He said the NTP uses a small number of slices from the blocks, although punches could theoretically be used.

Dr. Novak said he was unclear why the group was doing an antibody array for looking at the proteins as opposed to doing immunohistochemistry. Dr. Merrick said the decision was driven by the ability to work in a 96-well plate format, and there were drawbacks associated with the immunohistochemistry approach. Dr. Novak discussed validating the protein signature in comparison with the RNA, asking Dr. Merrick if it would surprise him if there were no necessary agreement between alterations in RNA and protein. Dr. Merrick agreed and said that some discordance between transcript and protein levels was a long-standing issue between the two ‘omics disciplines.

Dr. Novak asked if it was possible to pick up microRNAs with this method. Dr. Merrick said it was possible and was next on the team’s list to do. Dr. Tice added that one of the reasons NTP is interested in the qNPA methodology is that it does have a platform for looking at microRNAs from FFPE tissues.

XIX. Tox21 Activities: A Bioinformatics-based Approach to Identifying Assays that Query Human Health Effects

A. Presentation

Dr. Scott Auerbach, NIEHS/NTP, briefed the BSC on this Tox21-related activity, a nascent project that will be evolving over the next year or so. The goals of the program, he said, are directed at filling in the variables in the statements (1) disease Y hazard is queried most effectively by assay A, B, C… and (2) assay X queries disease hazard A, B, C…

Dr. Auerbach displayed a bidirectional flow chart from disease to assay to further illustrate the concept. Starting with disease there are several variables that can be considered, including pathways, biological process networks, intermediate phenotypes, tissues, organs, cell types, and genes, along with the many interactions and intersections of those areas. He said there are three ways to derive the outlined relationships, including literature mining, functional genomics, and genetics/genomics.

Dr. Auerbach described the approach being used for literature mining in the project. Specific steps include: (1) assembling a list of all genes in the human genome and associated annotations, (2) extracting disease-gene relationships from database resources, (3) mapping associated genes to the human gene reference list, (4) calculating a cumulative score for all genes for a given disease, and (5) determining assay development tractability/feasibility based on a "druggability" score.

Disease-gene relationships have been extracted from several databases, including the Comparative Toxicogenomics Database (CTD), CoPub, GeneCards, Entrez Gene, and Phenopedia, with its HuGE Navigator. Functional genomics resources include NextBio, the Unigene Body Atlas, GeneGo, and Ingenuity. Data resources for static genetics are NextBio and Phenopedia. He described each of the databases, and discussed how they could be mined to help achieve the objectives of this project.

Druggability is a critical concept in this activity. Identifying with what in the genome a chemical interacts is necessary to determine how to build a particular assay. Dr. Auerbach explained that the protein encoded by a gene is considered druggable if it contains a conserved protein domain that has been shown to bind to small molecules. There are six druggable genome sources, which are mined to give a specific gene a druggability score ranging from 0-6, using a binary scoring system.

He related some results to illustrate the process, first in type 1 diabetes. Showing results for several individual genes, they are scored using hand-curated or automated curation databases. For example, a particular gene associated with the disease under CTD, a hand-curated database, would receive a score of 2, because hand-curated databases are trusted more than one based on text mining only. If it appears in an automated curation database, it would receive a 1 score. The scores across the databases are summed and ranked, along with the gene's druggability score. Existing assays are also listed. This yields a list of prioritized assay targets. He cited a paper published in *Nature* (Ueda *et al.,* 2003) showing an association between the top-ranked gene on his list, *CTLA4*, with type 1 diabetes, illustrating the fact that this method does reliably identify genes that have been associated with a particular disease.

He showed another list of results for type 2 diabetes, along with a paper from *Nature Genetics* (Altshuler *et al.,* 2000) supporting the association of *PPARγ* polymorphisms with differential

susceptibility to type 2 diabetes. This illustrated that the approach can identify genes that have a causal relationship to specific diseases. He showed results of his obesity analysis, which was a bit more complicated given the fact that in the literature it is usually related to causal effects of other diseases. *PPARγ* was the top-ranked gene, likely due to its relationship with type 2 diabetes. Interestingly, a distribution of targets ranked highly, not just adipocyte targets, but neurological ones as well, affecting eating behaviors. Thus, targets can be determined in multiple tissues that relate to obesity.

In future work, Dr. Auerbach said he hopes to relate functional genomics of pre-disease states to assay selection, as they may yield more causality pathways, which may be good assay targets. He plans to consider target promiscuity, to identify targets where binding is likely to occur. He also intends to integrate gene-gene relationships to identify novel targets.

In summary, he noted that the BSB (1) is integrating multiple genomics/bioinformatics resources to identify genes and pathways associated with disease; (2) plans to identify or develop, based upon the associations, *in vitro* assays that will predict chemical hazard in a disease-centric fashion; and (3) anticipates that the findings will facilitate both prioritization of chemicals for focused *in vivo* studies and development of more cost-effective, efficient and informative integrated, human-centric testing strategies.

He mentioned that independent efforts in this research area are ongoing at NCGC, EPA, and have been published by a group from NIEHS. He also noted the upcoming NTP Workshop: The Role of Environmental Chemicals in the Development of Diabetes and Obesity, to be held in Raleigh, NC, January 11-13, 2011.

B. BSC Questions

Dr. Miller asked whether the overarching assumption is that druggable targets are a subset of toxicological targets. Dr. Auerbach said he believed that is the case for non-genotoxic chemicals that act by a pharmacological mechanism to produce toxicity.

Dr. Birnbaum asked Dr. Auerbach to elaborate on how he plans to look for early steps in the disease process. He cited the example of animal studies associating a high-fat diet with diabetes; that there are early changes prior to the initiation of disease. Describing other examples from his work, he said often those early changes are associated with risk of disease, whereas later changes are less likely to associate with risk.

Dr. Faustman suggested a slightly different approach; rather than curating the database with known disease pathways, she suggested looking for which early response gene pathways are networked with clinically diagnosable disease, or even look at what percentage of the databases are able to inform the pre-disease state. Second, she said she is not convinced about the value of druggability, that “some of our most notorious toxicants don’t know they’re supposed to be specific” and, thus, do not fit into druggability paradigms. So one gets low signal but multiple signals across the druggability space. In response to the first comment, Dr. Auerbach agreed that using a network-based approach focused on pre-disease is a goal of the project, but would be quite challenging with respect to identifying targets for assay development. In response to

the second comment, Dr. Auerbach acknowledged that using the current approach to identify assays that act through non-pharmacological actions (e.g., reactive chemicals) is a challenge and will need further consideration as the approach evolves. Dr. Auerbach went on to justify the initial approach that incorporates a consideration of druggability by indicating that the assay throughput at the NCGC is limited to 2 to 4 assays per month and hence the NTP would like to identify assays that produce robust, usable data in all cases. Assay methods to query traditional druggable targets are well developed and generally robust and therefore this is the reason the druggability criteria were used. Regarding the disease to genes pathway slide, Dr. Faustman suggested reordering the sequence of events to: disease > tissue > organ > early response > pathways > assay, i.e., tissue vs. non-tissue specificity should be considered earlier in the process.

XX. Concept Review: The Mouse Methylome Project

A. Presentation

Dr. Auerbach, substituting for Dr. Jef French, briefed the BSC about the Mouse Methylome Project, another activity central to the goals of Tox21. The Mouse Methylome Project's main goal is to sequence the methylome of three inbred mouse strains commonly used in toxicity studies. It is believed to be a launch point for incorporating genome-wide methylome assessments into toxicity studies.

Dr. Auerbach explained that the methylome comprises the positions of all methylated cytosines in the genome. It is one component of the epigenome, which controls non-genetic inheritance of phenotype. The biological role of the methylome is quite diverse; it is involved in X chromosome inactivation, imprinting, embryogenesis, gametogenesis, the establishment of cell type-specific patterns of gene expression, and silencing of repetitive DNA elements in healthy and diseased cells. In disease, it is known to be involved in cancer, through increased methylation of tumor suppressor genes and global demethylation, which leads to chromosomal instability. It is also involved in neurodevelopmental disorders, neurodegenerative and neurological disease, and autoimmune disease. The methylome varies as a function of genetics, sex, age, nutritional status, chemical exposure, and disease state.

Evidence also suggests that methylome variation contributes significantly to disease susceptibility. However, to date, no one has addressed methylome variation on a high resolution, genome-wide scale. The proposed project would create a high-resolution methylome map, and then would determine how the methylome varies as a function of genetic background, sex, and parental inheritance. Dr. Auerbach said this will contribute to a basic understanding of how the methylome influences disease susceptibility and will facilitate the incorporation of methylome assessment into toxicity testing.

The model system chosen was the liver in the B6C3F1 inbred mouse strain, and the two parental strains, C57BL/6N (B6) and C3H/HeN (C3). The liver is approximately 80-90% hepatocytes, making it relatively homogenous, avoiding the issue of cell type-specific methylome variations. The B6C3F1 strain has been used in NTP assays for 30 years, so there is much historical data. The liver is a common target organ in cancer bioassays. There is also

a spontaneous incidence rate of liver tumors in these mouse strains: <10% in B6, 100% in C3, and 30-50% in B6C3F1, with a higher rate in males. This increased liver tumor incidence often occurs following treatment with non-genotoxic chemicals, and it has been associated with alterations in the methylome.

The specific aims of the project are to: (1) create an in-depth, genome-wide map of the liver methylome of B6, C3, and B6C3F1/N mice; (2) identify regions of the liver methylome that vary within and between strains; (3) identify regions of the liver methylome that vary between sexes; (4) identify heritable regions of the methylome and how the heritability of these regions varies; and (5) correlate DNA methylation patterns with the transcriptome at a quantitative (expression) level and qualitative (splicing) level.

The animal study will include 10 breeding pairs: five B6 X C3 outcrosses, and five C3 X B6 outcrosses. Necropsies will be performed at 20 weeks, which will match with NTP archival materials relating to subchronic toxicity studies used to set the maximum tolerated dose. The liver and other homogenous tissues (brown and white adipose, cardiac muscle, skeletal muscle, brain, and sperm) will be collected from all parental mice and a male and female F_1 mouse from each breeding pair. DNA and RNA will be isolated from the liver, and other tissues will be archived for future use as warranted.

Several technologies will be used, all of which employ Next Generation Sequencing (rapid high throughput sequencing technologies employing short reads and alignment). First, genomic DNA sequencing will provide the sequences of each individual mouse genome to allow for subsequent mapping of BIS-Seq, ChiP-Seq, and RNA-Seq reads. The bisulfite-treated DNA or BIS-Seq read will provide a high-resolution map of the methylome. The sequencing of affinity-purified, methylated DNA (ChiP-Seq) will provide a lower-resolution map of the methylome, and is needed because future studies in NTP FFPE samples will use this method. RNA-Seq will sequence the transcriptome, providing a detailed map of RNA expression to be correlated with methylated cytosine sequences.

Dr. Auerbach provided considerable detail regarding sample analysis in the project, which has been broken into phases due to its resource-intensive nature. Phase 1 will focus on one quartet: a C57BL/6N female, a C3H/HeN male, and a male and female B6C3F1. They will undergo genomic DNA sequencing, BIS-Seq, targeted methylome re-sequencing of specific sites (ChiP-Seq), and whole transcriptome expression profiles with RNA-Seq. In Phase 2, BIS-Seq, ChiP-Seq, and RNA-Seq will be performed on a C3H/HeN female, a C57BL/6N male, and a male and female C3B6F1. In Phase 3, additional replicates and targeted re-sequencing will be performed as necessary to reach statistical conclusions in regard to within and between strain variations.

Describing the data analysis techniques planned for the project, he said bioinformatic and statistical methods to be used are still under development. First, the team wants to determine and catalog the genome-wide cytosine methylation patterns. Then, they will determine intra- and inter-strain variation in the methylome. Also, sexual dimorphisms in DNA methylation patterns will be determined, and local methylation patterns in the genome will be correlated with

quantitative and qualitative variation in the transcriptome. They will also identify heritable regions of the methylome.

Expected outcomes include a definitive map of the differentiated liver methylome and transcriptome, a definitive map of inherited genome methylation patterns (imprinting), and knowledge of the regions of the methylome that vary in genetically identical individuals, regions that vary as a function of genetics and sex, and the performance metrics of the two different methods for evaluating the methylome.

The project is ultimately expected to: (1) facilitate the identification of chemicals that perturb the methylome during different life stages, (2) assist in the identification of methylome-based biomarkers predictive of (hepato)carcinogenic hazard, (3) identify regions of the methylome that are particularly susceptible to environmental influence, and (4) assist in understanding the basis of individual susceptibility to disease.

In the future, the project will extend its approach to other tissues collected and to other inbred strains, if that is warranted. Also, the group plans to develop tools for methylome analysis of NTP FFPE archived tissues and archived human tissues.

B. BSC Questions

Dr. Birnbaum expressed strong support for the project and got clarification that the research would use five animals of each strain in each of the crosses, not one, and would be looking at a male and female progeny from each of the outcrosses. She asked if the ChiP-Seq approach would be adequate to pick up CpG shores, as opposed to islands. Dr. Auerbach said it would.

Mr. Janzen asked if the group had considered looking at histone methylation as part of the studies. Dr. Auerbach replied that due to the scale of the project, that had not yet been specifically considered, but it would certainly be a possibility later. He confirmed to Mr. Janzen that the NTP would preserve additional tissues for later studies.

Dr. Faustman noted that the dynamics of methylation patterns are being implicated for many disease conditions, and that life stage is an extremely important element. She said that although there was reference to life stage in one of the slides, there did not appear to be any plan for interim sacrifice to assess life stage outcomes. She felt that the dynamics of methylation patterns by life stage have been underestimated, and that the group has an opportunity to address that. Dr. Auerbach said it would be advantageous to have a number of interim sacrifices, but that assessment of life stages can be accomplished by going to the archival tissue and looking at time points of those interim sacrifices, as warranted by observation of variations. It would be a targeted approach addressing identified loci. Dr. Bucher added that the targeted approach would be used to develop probes for those regions that are most informative, allowing more efficient and more informative attention to the life stage questions.

Dr. Birnbaum noted that NIH is conducting a very large Roadmap project that is mapping the epigenome, including the methylome at a deep sequencing level for over 160 human cell types.

There are also centers looking at histone modifications and microRNAs. She said she shared interest with Dr. Faustman in looking at developmental questions, but that this project would be a starting point to creating the ability to address some of those questions. Drs. Birnbaum, Faustman, and Auerbach further discussed the choice of determining the methylome status at 20 weeks of age of the mice, in that it is not a developmental stage as such. Dr. Auerbach explained that the issue is the ability to age match, prior to the development of cancer, and complement the age of archived samples so that the chemical treatment would be the variable.

Ms. Rudel asked about the impact of pregnancy on the mouse liver, and whether the study should include an age-matched virgin mouse. Dr. Auerbach replied that the design includes non-pregnant female siblings of each strain in an attempt to control for that confounder.

Dr. Wiltshire commented that it was "an enormously valuable project," but cautioned that the difficulty of going back to archival tissues should not be underestimated, due to considerable variability. Dr. Auerbach said there were standard NTP protocols to deal with some of the issues raised by Dr. Wiltshire. Dr. French added that the NTP mouse strains are randomly assorted and assigned to treatment groups, to help minimize the impact of variability.

XXI. The Future of Tox21 at NTP

Dr. Tice summarized the Tox21 presentations and discussed the future of Tox21 at the NTP. He emphasized that Tox21 is a "community resource" project—a research project devised and implemented to create a set of data whose primary utility will be as a resource for the broad scientific community. Such projects have become increasingly important as drivers of progress in biomedical research.

Dr. Tice reviewed a timeline depicting the significant events in the formation and execution of Tox21, commenting that although it has only been formally running since 2008, "we've actually made, in some ways, a lot more progress than I anticipated."

He mentioned Tox21 outreach and interactions, citing nearly 200 presentations by Tox21 members since the original MOU in 2008. He also mentioned interactions with bodies from the European Commission, Health Canada, the U.S. Geological Survey, toxicology laboratories from the U.S. Department of Defense, and the Society of Toxicology.

Dr. Tice reviewed the Tox21 organizational chart again, noting that there are points of contact from each agency, the four working groups, and members from each agency, working together in a dynamic process that is a team effort. He also showed again the graphic depicting the concepts, working groups, and activities involved in Tox21, adding many arrows delineating the interrelationships among the various entities. He shared again the graphic depicting the specific tasks before Tox21, based on level of difficulty against time. He classified the many points by color to depict which tasks had been accomplished, which are still in progress, and which are scheduled through 2018.

He reviewed some of the other activities of the BSB, which had been mentioned briefly. They include the development of assays and informatics tools through the NIEHS SBIR/STTR program and efforts to link basic research to Tox21 via collaborations with NIEHS intramural

scientists. Other important collaborations were established by providing scientists with the NTP 1408 compound library and with samples of treated cells from the NCGC, and by working with scientists to further characterize qHTS results.

Dr. Tice displayed a graphic depicting the interactions between the BSB and several other groups within NTP, stating that Tox21 data will be used by NTP to: (1) prioritize compounds for more extensive toxicological testing and the kinds of toxicological studies that should be conducted and/or endpoints that should be evaluated; (2) interpret results obtained in laboratory animal toxicological studies; (3) investigate the relationship between genetics, environment, and disease; and (4) develop prediction models for animal and human disease.

He delineated for the BSC a number of limitations inherent to Tox21 and the questions and issues that need to be addressed and overcome, respectively, if this approach is to be scientifically accepted for its intended purpose. These include a recognition that not all compounds can be tested in HTS, nor are all *in vitro* assays useful for HTS. Furthermore, the Tox21 partners need to identify which assays are needed to cover the breadth of toxicity pathways, and which cell type(s) is the most appropriate for HTS assays that do not use reporter genes. For those that do, the partners need to identify the optimal format for reporter gene assays. Other issues include: (1) how human primary cells or differentiated stem cells could be used in HTS, (2) how xenobiotic metabolism can best be included in *cell-based* assays, (3) how to measure interactions between cells and tissues, (4) how to distinguish between active, inconclusive, and inactive compounds, (5) how to distinguish between statistical and biological significance, (6) how to extrapolate from *in vitro* concentration to *in vivo* dose, (7) how to evaluate for low dose, chronic effects *in vitro*, (8) how to distinguish between non-adverse and adverse perturbations, and (9) how to obtain the critically important human toxicity data.

Dr. Tice emphasized the importance of viewing Tox21 as a team effort, with participation by the international scientific community being an important element. The process must be science-driven, with clear milestones and quality assessment. He said the technology matters, in that new technologies are frequently created to meet new needs. He reiterated that it is critical to make all data publicly available, and noted that large scale data of this type is hypothesis-generating.

He said prioritization and prediction depends on comprehensive suites of *in vitro* and *in vivo* assays in a tiered approach, as is being pursued in Tox21. He added that targeted animal tests are needed to complement and demonstrate the relevance of *in vitro* tests. He said computational models of toxicity pathways are necessary. He concluded his remarks by mentioning that human toxicity data are essential if predictive models of human disease are to be demonstrated to be relevant.

A. Tox21 Activities – Mouse Methylome Project: BSC Discussion

Dr. Miller, first lead reviewer, said he felt the Mouse Methylome concept was clearly articulated, and the validity was well laid out. He rated the project's potential impact as moderate to high. He would have liked more detail about the aim of determining how the differences in cytosine

methylation will be linked to differential toxicities or disease outcomes, which he saw as "the crux of the whole matter." He encouraged the group to soon move on from mouse livers to diseases and organs that may have more relevance to human toxicities. He approved of the stated intentions to integrate with the epigenetics work related to the NTP archive and with Dr. Rusyn's work on inter-individual susceptibilities. Overall, he felt that the Tox21 mouse methylome project would provide a great database, and that the project was being greeted with much enthusiasm.

Dr. Cattley, second lead reviewer, said both stated aspects of the rationale appear valid. He felt that the overall significance of the project is difficult to assess given the many unknowns that exist at its beginning, but he rated the importance as moderate. He felt that the methylome is the right target of investigation, but questioned whether the mouse liver is the right site in which to do so. He wondered if it might be addressed in humans as opposed to the mouse. He also wondered if there might be mouse tissues other than liver that might give a more direct link to human disease and be more translational, given the lack of links between human and mouse liver cancers.

Dr. Faustman, third lead reviewer, said the more that can be learned about the methylome, the better to understand variability in response, so she was very supportive of the project concept. She felt that it might take a multi-stage process with more resources to adequately address the pertinent questions. She again expressed her opinion that the project needs to look across the life stages.

Dr. Jewell, fourth lead reviewer, said that as a biostatistician, Dr. Auerbach's statement that the statistical and bioinformatics tools for the project were still under development made him "a little nervous." He wondered if there was sufficient confidence in the ability to analyze the data to address the questions being explored.

Dr. Wiltshire, fifth lead reviewer, said he felt it was a very valuable, enormously important project that fits in well with the other Tox21 activities. He rated it as a high priority and felt it remains to be seen how well the endeavor to link the data to archival data will work out. He noted that it is a scientific project as opposed to a toxicological one, in that there is nothing toxicological being done. He would like to have seen even one chemical added in one lane, just to see what range of methylation status change might occur.

B. Tox21 Activities (in General): BSC Discussion

Dr. Miller said he was excited about the overall concept and supportive of the goals expressed, but did have some concerns and suggestions. He commended the presenters for acknowledging some of the limitations inherent in the proposals, but felt that in some cases the plans for addressing those limitations are not worked out as well as they might be. Some of the areas could be better integrated, although he recognized that some of those plans emerged in the discussions. For some of the proposals, he considered the performance criteria unclear, particularly criteria for when an unsuccessful effort might be halted.

Dr. Miller said that WormTox is “a world-class operation” and wondered when it might be applied to the rest of the Tox21 assays. He hoped that some of the tools developed for the transgenic worms might be applicable to some mammalian models. He said the methodology might be stronger as a problem-solving tool and a mechanistic tool than as a screening tool.

On the inter-individual susceptibility, Dr. Miller said the appropriateness and adequacy of the program as it relates to the Tox21 program are acceptable. He approved of Dr. Rusyn’s recognition of the limitations of his system, and encouraged him to continue to face those limitations as follow-up studies with additional chemicals, in hope that some of the potential variables could ultimately be filtered out.

Dr. Miller expressed enthusiasm about the NTP archive project, with the massive amounts of data that can be leveraged. He was concerned; however, that the earliest time points likely to have been used in most of the archival studies may be too late to identify key interactions. He felt that microRNAs were worthy of further exploration. In terms of how to define which ones to study, he suggested starting with plasma samples, which could represent a translational opportunity. He supported linking the archival activities with some of the other Tox21 activities, particularly the Mouse Methylome Project. He wondered how many FNs could be tolerated, in terms of performance criteria.

Regarding the bioinformatics approach to disease pathways, he suggested there may be an opportunity there to focus more on chemical signatures rather than disease signatures. He reiterated his earlier comment that he doubts that the druggable world encompasses all of the toxicological targets.

Dr. Miller closed his review by expressing that he was “extremely excited” about the current direction of NTP as embodied in these efforts.

Dr. Cattley lauded the WormTox project, stating it would be interesting to see how hepatic, renal, and respiratory system toxicants would perform in the assay. He also suggested future assessments of locomotion in the worms, and potential effects on the nervous system.

He was enthusiastic about Dr. Rusyn’s program, but suggested consideration of using human hepatocytes as a model system.

Regarding the NTP archive data-mining project, he agreed that the timing of sacrifice was likely late in many of the studies, in terms of characterization of toxicity, potentially resulting in detection of many secondary responses. He added his concern that the tissues were not originally collected for the purpose of nucleic acid analysis, leading to variation in how the samples were collected and processed.

Regarding the bioinformatics initiative, he said the program seems nascent and it was unclear whether the target genes for the diseases are tissue-specific with respect to their relevance. He felt that it was a very complex undertaking, but that it would be important to be able to determine whether particular genes are causative of disease, or adaptive.

Generally, he felt that these Tox21 initiatives focus on important aspects of the movement in toxicology to address hazard assessment in new and effective ways. He wanted to see more emphasis on the key question of predictability of endpoints, along with an earlier emphasis on translational research, starting now, with studies to understand variability in biomarkers. He felt that the methylome project emphasizes one aspect of gene regulation, but that consideration of microRNAs and other RNAs that regulate gene function needs to be elevated.

Dr. Faustman said the NIEHS needs to be conducting the Tox21 activities, and that the portfolio is "exciting, interesting, and on target for the types of questions that come up under Tox21." She was concerned that perhaps in some of the activities there was an imbalance between scientific expertise and mechanistic understanding, and the need to get the studies done quickly. She asked for a bit more articulation of the approach to those issues. The approach to prioritization of compounds was not intuitively clear to her. She was pleased that limitations of the program had been explicated, and felt that they should have been presented for each one of the activities. She recommended that the program pull more expertise from the outer communities, getting other areas of expertise (e.g., teratology) engaged and involved, in that the program seems "cancer-centric" presently. As she put it, "you should be on the road for Tox21…and these resources."

Dr. Jewell said Tox21 is a "treat to statisticians to have all these data and clearly important scientific questions." He wondered at times whether or not the data would be interesting or relevant. He recommended including attention in each presentation to the major statistical challenges being faced by the investigators. He mentioned an earlier comment that it may take days and weeks to collect the data, but it would take years to analyze it, noting that it brought up a real and important issue.

Dr. Zelikoff considered the Mouse Methylome Project "extremely worth pursuing," but expressed concern that when conveying information to some of the stakeholders, as well as the lay public, it may be difficult to translate the methylome to a clearly linked chemical hazard, a change in phenotype, or a disease state. Regarding WormTox, she felt that it could not replace testing in animals, and that its results would still need to be validated in animals. She was concerned about how to translate and select doses, and the relevancy of doses for non-species extrapolation. She felt that the whole Tox21system is lacking in tests of mixtures. Regarding Dr. Rusyn's program, she said that use of the immortalized lymphoblastoid cell line is problematic in that any immortalized cell line loses or gains some properties that bring it further away from the human genotype, which is ultimately what is being examined. For the archives program, she said that care should be taken as to sample purity, particularly in light of sample degradation.

Dr. Wiltshire said he was not clear how WormTox would reduce the use of animals, in that it was unclear how a feeding or movement or reproductive assay would be assessed in terms of a rat cancer phenotype, with there being no correlation between the two at the moment. Generally regarding Tox21, he saw the Tox21 activities as focusing almost exclusively on acute exposures, which is partly a result of the technology that has been developed to look at toxicity.

He said that at some point the program should go back and take another look at other technologies or methods that would allow investigation of low-dose, long-term exposures.

XXII. Report of the NTP Associate Director

Dr. Bucher awarded a certificate of appreciation to retiring BSC member Dr. Bunton.

XXIII. NTP Testing Program Research Concepts: Overview

Dr. Scott Masten, Director of the NIEHS/NTP Office of Nomination and Selection, briefed the BSC on the nomination process and its charge. He said the NTP studies individual or classes of substances judged to be possible hazards to public health or for which toxicological knowledge gaps exist. He noted that issue-based nominations and hypothesis-driven research are also considered. He illustrated the breadth of the NTP's research portfolio, mentioning the many areas of emphasis in the testing program, highlighting how the substances to be presented in the meeting fell into those major thematic areas.

Dr. Masten briefly reviewed the process for development of NTP research projects. Most are developed in response to external and NIEHS/NTP nominations, which are subjected to multiple levels of review, with not all nominations leading to a research program. Studies include both substances and issues, and an iterative approach to study design, conduct, and analysis is used.

Dr. Masten used a flow chart slide to illustrate the NTP study nomination process. He noted that nominations can and do come from a wide variety of sources. There has always been an interagency review step in the process, but in the past nine months a new process had been instituted, utilizing specific points of contact at the other Federal agencies to coordinate their input on nominations and draft concepts. The projects to be presented at this meeting are the first to have undergone the new process.

In the sessions at this meeting, first the research concepts would be presented by the NTP project leaders, followed by public comments, comments from assigned BSC reviewers, and general BSC discussion. Dr. Masten said a research concept is designed to be a brief, informative document, outlining the scope of the project, but is not intended to be a full study design. He briefly reviewed each of the research concepts to be presented.

He also summarized the BSC's charge related to its review of the proposed research projects. The BSC is to review and comment on draft research concepts and determine whether the proposed research projects are an appropriate use of NTP testing program resources. The BSC is to comment on the clarity and validity of the rationale for the program, comment on the merit of the program as it relates to the mission and goals of the NTP, rate the overall significance and public health impact of the program, comment on the scope of the proposed program, and provide any other pertinent comments.

XXIV. NTP Testing Program Research Concept: Exposure Characterization and Reproductive Health of Men Working with Bisphenol A in the United States

A. Presentation

Dr. Steven Schrader, NIOSH/NTP, presented the research concept (co-led by Cherie Estill) to the BSC. The nomination emerged from the recent publication of papers showing a correlation between Chinese occupational BPA exposure and male reproductive health. The program would address the question of whether U.S. exposures are similar, and if so, will the same findings be seen in U.S. workers? Overall, the goal would be to characterize exposure and evaluate the reproductive health of U.S. men who work with BPA.

Dr. Schrader provided some background related to BPA; its uses, the virtually ubiquitous exposure of the U.S. population, and the fact that animal toxicological studies are mixed on the association between BPA and adverse effects on reproductive function. He noted that the 2008 report by the NTP's Center for the Evaluation of Risks to Human Reproduction (CERHR) had recommended human occupational exposure assessment to clarify BPA exposures and internal doses in workers, urging studies of the effects of adult exposures on reproduction.

Dr. Schrader reviewed two recent papers reporting on BPA exposures and reproductive markers or outcomes. The first (Meeker, *et al*., 2010) showed that as BPA urine concentration increased in a general population, follicle stimulating hormone (FSH) levels increased, inhibin B decreased, and the estrogen-testosterone ratio decreased. A similar study (Mendiola, *et al*., 2010) in fertile couples reported that as BPA urine concentration increased, the free androgen index decreased; however, no effect on semen quality was detected.

Dr. Schrader provided more details about the Chinese occupational studies, which reported that BPA-exposed workers have adverse sexual function and that BPA urinary concentration increased as sexual function decreased, in both environmental and occupational populations. The group's most recent paper (Li, *et al*., 2010) looked at semen quality as per BPA urinary concentration, and found that as concentration increases, sperm count, viability, and motility decreased.

Dr. Schrader identified the key issues involved in the proposed research: (1) occupational exposure levels to BPA in the United States are unknown, (2) Chinese occupational exposures have been associated with decreased semen quality and sexual function, (3) BPA induces adverse reproductive effects in male rats and mice, and (4) U.S. environmental BPA exposures have been associated with endocrine changes in adult males. The proposed research program is intended to: (1) identify industries, processes, and worker populations with potential exposure to BPA; (2) evaluate occupational exposure to BPA by performing environmental and biological monitoring; and (3) determine the reproductive health of the male workers by assessing their reproductive endocrine profiles, semen quality, and sexual function.

The industries with likely exposure may include producers of polycarbonate plastics, epoxy resins, and thermal paper, as well as foundries. The study will determine which worker populations have been exposed, and their exposure levels along with job classifications. As part of the effort to evaluate occupational BPA exposure, the researchers will conduct environmental monitoring using air and surface sampling methods employing LC/MS and UV methods. Biological monitoring will also be performed by taking pre- and post-shift urine samples on the first and second days of the workweek. Dr. Schrader provided details of the

planned methods to determine the male workers' reproductive endocrine profiles, semen quality, and sexual function.

He summarized by reiterating that the program is intended to characterize and document BPA exposure levels in the U.S. workplace, and to determine if U.S. occupational exposures are associated with adverse reproductive health in male workers.

B. BSC Questions

Dr. Schrader confirmed to Dr. Birnbaum that females would be included in the initial walk-through studies. Dr. Birnbaum suggested incorporating them in the overall study design. Dr. Schrader pointed out that that would make it a much larger study. He said NIOSH would certainly be capable of doing so, in that there is a group within the agency that specializes in female reproductive health. He elaborated that the researchers would have a better handle on the question once they start doing their initial walk-throughs. Once a company to be studied is identified, NIOSH industrial hygienists perform a walk-through of the facility, getting an initial idea of where exposures might be taking place and the locations of worksites they would return to for the complete study. He indicated that if populations of women working with BPA were observed, NIOSH would submit a proposal to study women as well.

Dr. Schrader responded to Dr. Faustman that there is no current occupational standard for BPA and no current public alert regarding BPA. Returning to the issue of women and BPA, Dr. Birnbaum stressed that it was an important issue to consider, particularly since there had been reports of effects in children associated with their mothers' first-term exposure. Dr. Schrader said that traditionally in manufacturing, the number of women is quite low, and the issue is one of having ample population to assess. Dr. Birnbaum noted that NTP is preparing to conduct a study of BPA exposure in cashiers that handle thermal printer paper (known to have high BPA levels), with the anticipation that a large number of that population would be female. She asked whether that might be a population that NIOSH might pursue, if BPA levels were found to be elevated. Dr. Schrader confirmed that NIOSH would be interested in that cohort.

Dr. Loomis asked for elaboration on the number of exposure measurements NIOSH would perform. Ms. Estill responded that for aerosol, they would measure two full days of exposure, by personal monitoring. Surface samples would also number four per person over the two days, as would urine samples. Dr. Loomis asked how many people would be involved. Dr. Estill said there would be a minimum of 10 people per facility for the walk-throughs, and that the reproductive study would enroll a total of 150 people—30 per plant at 5 plants being the goal. The subjects would be categorized by low, medium, or high exposure, based on their job classifications or locations. Dr. Schrader added that typically a wide distribution of exposure is seen in workplace settings.

Dr. Paul Howard, FDA, noted that the NHANES data for BPA did not exhibit spikes in terms of population exposures. He said that for the most part, BPA production does not take place in "mom and pop shops;" but rather in large industrial plants. He asked if smaller operations making plastics would be part of the target of this project. Dr. Schrader said that would be very difficult to do given the difficulty of getting enough statistical power with a very small group of

workers. Dr. Howard said his only concern is that large industries tend to be more chemical hygiene-oriented than small operations. Dr. Estill noted that NIOSH has a Health Hazard Evaluation Program, which could provide services to smaller plants to reduce exposures.

Dr. Faustman asked if NIOSH was planning to assess the recycling industry. Ms. Estill said they were planning to look at thermal paper recycling, but would consider other forms of recycling as well. Ms. Rudel questioned if the data being collected could be used to develop an exposure measure to specifically measure inhalation and dermal exposures. Ms. Estill agreed that that would be good. She said NIOSH would see if the aerosol and surface samples correlated with urine measures. Ms. Rudel asked why a surface wipe was being used rather than a skin wipe. Ms. Estill said she had not found a hand wipe as of yet, but may do so as the program progresses.

C. Public Comments

Mr. Joseph Manuppello, Research Associate at People for the Ethical Treatment of Animals (PETA), said this study represented an approach that PETA would like to see NTP take more frequently in the future. He asked for comments from the BSC about how the data generated in this and other human studies might reduce "the very large number of animals being used" at the FDA and other government agencies.

D. BSC Discussion

Dr. Faustman, first lead reviewer, noted that a significant amount of information had been added in an addendum to the written proposal since she completed her written review. She said that this is a very visible and important area that needs more research and clarification, and that there is surprisingly little information available about occupational exposures to BPA. She thought the proposed project would begin to address the missing information, and although very supportive of the project area, suggested a need for additional clarification for it to move forward. She noted the clarifications on exposure data included in the addendum, but wondered if anyone had looked backward to estimate how high or how variable the exposures might have been. She recommended running a kinetic model, particularly since BPA's half-life is so short, since those results may affect some of the study's design issues. Some of the animal studies of BPA show that single-dose, high peak exposures are causing physiologic changes, and so such exposures may be important in an occupational setting, depending on how BPA is used. Dr. Schrader responded that the walk-throughs would help to answer some of the questions regarding sampling strategies. Dr. Faustman asked if a decision tree had been prepared to guide the project, e.g., in the event that only a very low level of BPA exposure was found in the workers, and how would that trigger a completely different approach. She expressed concern about the lack of details in the proposal regarding the study's endpoints, e.g., cycle status would influence several of the endpoints to be measured.

Dr. Faustman expressed concern about the fact that since public action had been taken on BPA, rightly or wrongly, there may be a need to issue an alert to the workers that exposure to BPA represents a potential problem. Dr. Toraason responded, stating that NIOSH does put out flyers to the affected industries for these types of studies, and that the industry uses the

informed consent process. Dr. Faustman said she was not concerned so much about the issue of human subjects, but worried about the idea of informing a small number of industrial entities that there may be a problem, without alerting the industry as a whole. She noted that BPA is a highly visible issue, with several countries having taken dramatic steps, including bans. She felt that a statement from OSHA, or another government agency, was needed. Dr. Toraason said NIOSH doesn't know whether an alert is warranted, and that is why the work is being done. Dr. Faustman said there were data available and that she did not want to see the NTP working on a problem where the broader industry has not been adequately informed that there is a concern. Dr. Toraason questioned what information would be included in the alert. Dr. Birnbaum added that at this point it is unknown whether workers in the United States have any more exposure to BPA than the general population, and that China does not have the same occupational standards as here. She suggested the first step would be to establish whether there is exposure, and then, if so, to communicate that information.

Dr. Jewell, second lead reviewer, said he felt that in terms of clarity and validity, the rationale for the proposed study was well-justified and well-presented. He agreed that the study should be extended to females as well as males. He said it was important to do this work given the current "frenzy" associated with BPA. He mentioned that it was unclear in the proposal how routes of exposure effects would be differentiated, particularly across different industries. He felt the study is strongly relevant and rated its public health impact as "high." He was critical of the lack of detail in the proposal, and was concerned about the scale of the study, particularly in how the sample size of 150 was determined. He strongly disagreed with the plan for regression analysis using 10 independent variables at baseline. He asked for more information about the outcomes that would be measured, and the primary outcome, noting that multiple outcomes would be a "fishing expedition." He felt the rationale for the sample size of 150 was incomplete, and that perhaps there should be two or three times that many subjects. He said the companies to be sampled should be determined soon, as there is concern that companies or individuals volunteering for the study could allow bias to arise. He expressed concern about the time of day the urine samples would be taken, whether diurnal variation in urine BPA concentrations had been taken into account, and suggested a small pre-study to address that question. He requested more information about potential confounding factors (e.g., obesity) and how they would be measured, as well as how occupational and casual exposures would be differentiated. He said he would need to see much more detail overall about the study before he would be confident that it would be of great value and add important information.

Dr. Masten responded, stressing that the proposal was a high-level approach to the problem, that many of the design details would be worked out in the future, and that the comments made would certainly be of value in that process. Dr. Schrader concurred, noting that NIOSH would work with their statisticians to ensure that the concerns expressed would be addressed. In terms of the selection of industries, he emphasized that NIOSH does that regularly, following a process using its legal right of entry to investigate health questions.

Dr. Teeguarden noted that BPA is now "a high-stakes game," with the exposure side being most important right now, particularly occupational exposure. In that context, he said, this study is "unquestionably of high value" to fill in a major data gap. Another issue to be considered is the

presence in the blood of free, non-conjugated, bioactive BPA metabolites and the conjugated form. This ties in with the question about route of exposure, as some routes may bypass the gastrointestinal tract. He said that although the proposed study did not address that issue, it should be considered for inclusion by taking some extra samples for that purpose.

Dr. Teeguarden continued by addressing several potential confounders to the study. First, internal versus external exposure: he said external exposure using wipes is valuable to help identify sources of exposure, but is the least valuable for trying to make unconfounded relationships between any endpoint and exposure. Thus, he recommended that the paradigm be more focused on internal exposure, particularly blood, which would allow measurement of the free versus conjugated metabolites. Since the implicit assumption in the study is that the workers would have higher internal exposures than the general population, he also recommended urine biomonitoring, preferably on a 24-hour basis, as a more definitive way to assess exposures than the use of wipes measuring external sources. He also noted that the ingestion of food is a major confounder, as it is known to be the major source of BPA exposure, and blood and urine concentrations spike following meals. Thus, measuring urine or blood levels pre-shift would actually be measuring peak levels following breakfast, and would confound the use of those measurements as a control or baseline to assess workplace exposures. He suggested finding a way to remove ingestion as a major confounder and being very aware of contamination as a confounding issue. He offered to share data he has regarding BPA blood levels following meals.

Dr. Schrader said NIOSH would be very interested to see that data, and that they were planning to take urine samples during the initial walk-throughs. He added that the reason for the air and wipe samples is to establish where some of the exposures are coming from as they go from facility to facility.

Dr. Novak concluded the session, stating that his sense of the discussion was that everyone considers the study to be a high priority, recognizing the caveats concerning the addition of women and careful consideration of study design elements.

NOTE: Dr. Wiltshire departed the meeting as this session began, and Dr. Faustman was absent for this session.

XXV. NTP Testing Program Research Concept: Cholesterol and Lipid Modulating Agents: Toxicological Approaches to Assessing Complex Mixtures

A. Presentation

Dr. Barry McIntyre, NIEHS/NTP, explained to the BSC that this nomination, from a member of the public, was actually in the form of two separate nominations: “Drinking water disinfection by-products: 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase inhibition and developmental toxicity” and “Drinking water disinfection by-products: interactive effects of antilipidemic agents and drinking water contaminants in producing developmental toxicity.” Given the obvious interrelationships of the nominations, it was decided to present them as one concept.

Dr. McIntyre said the concept would address two complementary toxicological aspects: (1) testing, improving, and validating approaches for cumulative risk assessment and (2) characterizing the toxicological outcomes of *in utero* exposure to mixtures of agents that affect lipid and cholesterol utilization in the fetus/neonate.

Noting that the NTP does not conduct risk assessments, he said that risk assessment is typically carried out on a chemical-by-chemical basis, but of course in real life we are exposed to complex mixtures of chemicals. He added that there are chemicals to which we are exposed that are not monitored for regulatory purposes (e.g., pharmaceuticals). The field of mixtures toxicology is emerging as an area of scientific and regulatory focus. Studies of mixtures with endocrine active compounds have demonstrated that chemicals that target a common signaling pathway/tissue can contribute to dose additive toxicity. He added that chemicals present at concentrations below their respective no observed adverse effect level (NOAEL) can incrementally add to a "total mixture dose" that could potentially elicit toxicity.

He illustrated the concepts of response addition, in which combining chemicals at the NOAEL levels would produce no adverse effects, and dose addition, in which all chemicals that target a common signaling pathway contribute to dose additive toxicity, even if individually they are at NOAEL levels. He also showed data illustrating that the dose additive model closely resembles the cumulative observed effects of several different chemicals that have different modes of action on a common tissue. These data demonstrate that it might be possible to use the dose addition model for other agents that target a common signaling pathway, tissue, or mode of action.

As an illustration of mixtures occurring in the environment, Dr. McIntyre showed environmentally relevant chemicals and pharmaceuticals found in wastewater effluent. The agents included phthalates, fibrates, statins, and others. Although they were found at very low concentrations, in some areas they may be present in many-fold higher concentrations, which is a growing concern in regions where water re-use is common. He added that the various agents could also be considered as models based upon their modes of action; phthalates being PPAR agonists, for example. The examples he showed had similar modes or complementary modes of action, illustrating the concept that although the agents themselves may be quite different, they may adversely affect the same tissue or organ system. The same concepts hold true for the effects of lipid modulating agents on fetal development—the agents have similar, overlapping effects, on a class basis.

These concepts generate the study's hypotheses: (1) collective modulation of cholesterol and fatty acid levels by cholesterol and/or lipid modulating agents that have complementary modes of action will result in similar (overlapping) adverse responses and (2) these responses should be able to be predicted by using dose-addition, response-addition, or integrated-addition models. Dr. McIntyre provided the proposed study's specific aims: (1) characterize the dose-response relationships for prenatal toxicity using select members of each class; (2) use data from Aim 1 to make toxicity predictions of the mixtures using dose-addition, response-addition, and integrated-addition models; and (3) conduct multiple mixture studies, both within a class

(i.e., 3 or 4) and across classes (e.g., 1 to 2 from each class; multiple classes), to test the predictions made in Aim 2.

The preliminary study plan is to review the relevant data in the public domain, to conduct limited dose duration perinatal toxicity studies (likely the most sensitive time of development) in the Harlan Sprague Dawley rat, to use the data emerging from those studies to make toxicity predictions based on dose-addition, response-addition, and integrated-addition models, and to conduct combination mixture studies to validate those predictions. He showed a matrix illustrating a variety of potential mixture studies, including various combinations of statins, fibrates, phthalates, haloacetates, and perfluoroalkyl acids (PFAAs).

Describing the significance and expected outcomes related to the studies, he noted that they would broaden understanding of the potential hazards of concomitant exposures, that the data would expand the knowledge base on mixture toxicity while exploring the utility and predictivity of the models, and that a better understanding of these toxicological challenges would aid in the design of future mixture studies.

B. BSC Questions

Dr. Birnbaum noted that this was a very ambitious range of studies and urged consideration beyond the immediate birth or developmental effects that might be seen, in that some of the effects might not be evident until the offspring are 90 days or older. She said there is much literature about "something from nothing" responses and she suggested using initial short-term, *in vitro* screening approaches to lessen the need for large-scale animal testing. She speculated that when the literature review is completed, perhaps the project would not require the breadth and depth currently being presented.

Dr. Sherley asked if there was a plan to include combination drug treatments in the literature search, and whether that type of information might inform planned modeling. Dr. McIntyre replied that some of the statins are co-administered, and that there are some limited toxicity data available.

Dr. Miller asked about the plan for measuring internal dose; whether that would involve systemic exposure, fetal exposure, or tissue exposure. Dr. McIntyre said there would be a comprehensive assessment involving all of those measures, as well as several others.

C. Public Comments

Mr. Manuppello related PETA's comments, pointing out that the concept is particularly vague regarding what chemicals or combinations would be tested, or which developmental endpoints would be assessed. He said that although a number of non-animal approaches are clearly suited to a preliminary investigation, none are mentioned or considered in the proposal. He cited several specific concerns about the scientific premises of the proposal, particularly the relevance to humans of developmental toxicity data in animals in the various compounds included in the testing regimen. He said the research concept "clearly fails to justify the use of animals that would be required for its implementation," and he called upon the BSC to reject it.

He also objected to the anonymous nomination, urging the BSC to demand more transparency in the nomination process.

D. BSC Discussion

Dr. Miller, first lead reviewer, said the overall concept presented seemed sound, but that the model needed to be clarified. Given the experimental approach described, he wondered if it was proof-of-principle as opposed to an attempt to recreate real-world exposure scenarios. He felt that the proposal was in-line with the NTP's mission, and praised the group for taking on a very challenging area of toxicological research. He ranked the proposal as moderate in terms of overall significance and public health impact.

Dr. Carney, second lead reviewer, said he looked to NTP to take the lead in mixtures research, an important unresolved issue. He deemed the concept proposal valid, and said Dr. McIntyre's presentation helped its clarity. He approved of the goal of moving an approach forward for characterizing mixtures, but wondered about the value of the second goal, characterizing specific responses. He advocated "re-spinning" the concept to elicit more testable hypotheses, comparing it presently to a "fishing trip." He felt that much of what was being sought in the project in terms of biological effects was already established without the need for another large project. With the infinite number of mixtures to which people might be exposed, novel approaches are needed to determine which ones to work on. He said a more systematic approach should be used to decide which combinations are important, perhaps using exposure data. A proof-of-concept for a workable approach to testing the concept needs to be cost effective, efficient in terms of animal use, time, and expense, and that jumping immediately into a pregnant animal model might not be a practical place to start. He advocated starting with a more tiered approach, similar to the ideas put forth in Tox21.

Dr. McIntyre responded stating that one advantage of using the proposed model is that it has a lot of power, with a great deal of data generated in a population believed to be more sensitive to effects. The initial studies can be used to validate the models, and then go to the literature, using fewer animals, and have some degree of confidence in the predictivity of the models.

Dr. Miller questioned the meaning of the term "validating the model," i.e., whether it meant within this class against these receptors, or a more generalized conclusion. Dr. McIntyre replied that that was an important question that was still being considered within NTP; these studies would clarify the utility of the respective model.

Regarding exposures, Dr. Carney noted that one could assume a worst case scenario, that in a mixture of even 50 compounds, in the dose addition model, a pharmaceutical would probably totally dominate the mix in terms of effects and dosing level, in that the other compounds would probably be at extremely low levels. He said that could just be assumed, and did not need to be proven.

Dr. Minor asked whether the group could just do a simple experiment with a statin in a cell, examining biosynthesis at different concentrations. Dr. McIntyre replied that the challenge is to determine modes of action, and that Dr. Minor's suggestion would be taken under advisement.

Dr. Birnbaum said the proposal needed more internal discussion, as well as additional discussion with the BSC. Regarding the potential domination of a pharmaceutical in the mix, she assumed that the researchers would use a ratio designed to mirror the one likely to be found in drinking water—not the dosages as such, which would be amplified in the experiments, but the ratio itself. She thought that approach would have some validity; however, it might not be applicable to young, pregnant women. She noted that with the Tox21 approaches presented at the meeting, there were many opportunities to test many environmentally relevant mixtures, including pharmacological agents, environmental agents, and food constituents. Such studies could be used to guide a more limited number of animal studies.

Dr. Carney urged that the approach be made simpler, since it is not intended to be an actual risk assessment, and the interactions would be occurring upstream of the fetus.

Dr. Novak ascertained the BSC had moderate enthusiasm for the concept and felt the need for significant refinement of the overall concept.

XXVI. NTP Testing Program Research Concept: N-Butylbenzenesulfonamide (NBBS)

A. Presentation

Dr. Cynthia Rider, NIEHS/NTP, presented the research concept for NBBS, a plasticizer found in polyacetals, polycarbonates, polysulfones, and polyamides. It is a high production volume chemical, which has been demonstrated to leach from polyamide cooking utensils. It has been detected in water samples from New Jersey and California, and in 1 of 42 adipose breast tissue samples.

NBBS was nominated by NIEHS for comprehensive toxicological testing based on its high production and use, with likely widespread exposure. There are limited toxicity data, and there have been structural alerts from QSAR analysis, along with some positive findings in toxicity studies. There are limited data available for toxicokinetics (TK) and absorption, distribution, metabolism and elimination (ADME), some data for reproductive toxicity and neurotoxicity. No studies have been done addressing chronic toxicity, carcinogenicity, or immunotoxicity. The one TK/ADME study showed rapid uptake and distribution into tissue, including brain, with little accumulation and rapid triphasic elimination. The one *in vitro* metabolism study available showed that hydroxyl metabolites were detected, but no conjugation of metabolites was noted.

Dr. Rider presented a summary of some of the short-term and subchronic studies that have been performed with NBBS. They showed that frequent targets of the compound are the nervous system, blood, and the male reproductive system. Standard tests have shown NBBS to be non-mutagenic. The manufacturer has tested NBBS in a reproduction and developmental toxicity screen using Wistar rats dosed by oral gavage two weeks prior to mating. Various adverse reproductive effects were seen in the parental animals and the F_1 animals. NBBS has also been tested in some HTS assays by NCGC, but it has not been tested with the complete battery of assays. It was generally not very active in the HTS tests, and did not elicit cytotoxicity or DNA damage. It did not act as an agonist or antagonist to androgen or estrogen receptors. It was a weak agonist of retinoid X receptor and weakly activated CYP2C19. However, only the

parent compound was assessed in the HTS tests. There are considerable deficiencies in the existing studies, and significant data gaps remain regarding immunotoxicity, neurotoxicity, and reproductive/developmental effects.

The proposed studies would consist of three phases. Phase I would be comprised of TK studies in rat and mouse via oral, dermal, and intravenous administration, which would allow for comparison of absorption from relevant routes of exposure. Also, *in vivo* toxicity would be assessed in Phase I via a rat perinatal dose range-finding study and an adult mouse oral 14-day toxicity study. In Phase II, there would be further *in vivo* toxicity studies, including a rat subchronic study with perinatal exposure and an adult mouse 90-day toxicity study. Phase III would be 2-year studies in rats and mice.

The studies are intended to: (1) fill-in large data gaps on a widely-used compound with high exposure probability and indications of neurological, developmental, and reproductive toxicity; (2) provide dose-response data to inform risk assessment; (3) provide dosimetry data to link environmental levels to toxicity in rodents; and (4) provide information for QSAR because structurally related compounds are also not well-studied, but are also common building blocks for manufactured compounds

B. BSC Questions

Noting that NBBS is a high production volume compound with a high potential for widespread general population exposure, Dr. Birnbaum wondered whether there was also high potential for occupational exposure, and if so, whether NIOSH had investigated it in worker populations. Dr. Toraason replied that although NIOSH did not know much about the compound, they would be interested in adding it to their list for exposure assessment. Dr. Birnbaum said it is becoming more and more evident that there are opportunities to investigate some of these compounds in human populations, and that if it is discovered that there is significant occupational exposure, there is an increased need for NIOSH/NTP to conduct experimental studies on them. Ms. Estill agreed that NBBS should be assessed by NIOSH. Dr. Howard concurred, particularly since NBBS is contained in some personal care products, noting that an occupational exposure study could contribute maximum exposure levels, which could help inform potential general population exposures. Dr. Birnbaum said she hoped NTP could overcome the difficulties with perinatal studies of exposure in mice, perhaps by using another strain. She said such studies are vitally important given that development is such a critical window of susceptibility, and considering the power of conducting studies in mice.

Dr. Loomis mentioned that it would be very useful to have an update of the National Occupational Exposure Survey, which was last conducted 30 years ago, to help provide data on occupational exposures to NBBS and many other compounds.

Dr. Bunton asked about the 28-day rat study Dr. Rider had mentioned, which showed hematologic effects, but for which details were not available. Dr. Rider explained that the study had been published in Russia in 1979, but the article had never been translated and had only been summarized.

Dr. Miller asked if NBBS had been studied in *C. elegans*. Dr. Rider said no. Dr. Miller noted that it might be a good candidate for WormTox, given its potential reproductive, developmental, and neurologic effects.

Dr. Faustman noted that NBBS is representative of a whole class of compounds for which there is very little information. She suggested that the pharmacokinetic, ADME, and *in vitro* characteristics should be focused on up front. It being a high production volume (HPV) chemical, she wondered whether anyone had submitted data on it for Registration, Evaluation, Authorization, and Restriction of Chemical Substances (REACH) or other HPV-related projects. Dr. Masten noted that the project did not have a specific class in mind, and that the sulfonamide substructure is very common. He said NBBS had risen in priority for consideration due to its HPV and use in many consumer products.

C. Public Comments

Mr. Manuppello provided comments from PETA regarding the research concept. He noted that Dr. Rider had already addressed several of the points he wanted to emphasize, but he did want to comment given that the proposed research is a large-scale animal testing program. He urged the researchers to acquire the full version of the OECD study Dr. Rider had cited, so as to avoid the potential for any duplicate testing. He also suggested use of an OECD Extended One Generation Reproductive Toxicity Study, which uses half as many animals as individual studies would. He said the need to conduct these studies in both rats and mice had not been justified. He commended NTP's consideration of *in silico* data, and said cell transformation studies should be carried out *in vitro* before any *in vivo* studies are considered.

D. BSC Discussion

Dr. Rudel, first lead reviewer, felt the rationale for the project was clearly stated and compelling. Based on that rationale, she said, the proposed research fits clearly with the mission of the NTP, with high potential significance and impact. She suggested expanding the scope of the research to add to human exposure data on the compound. She also suggested development of biomarker methods for the ADME work, which could then be used in humans. She was intrigued by the idea that NBBS might be representative of a class, and asked for more information about what other chemicals might be included and what their common properties might be. She thought that perhaps there should be additional genetic toxicity assays included in the project's scope. After testing is completed, the findings should be rolled into the evaluations of the HTS assays and QSAR approaches. She approved of the inclusion of those methods in the proposal, since that had not been seen previously.

Dr. Cattley, second lead reviewer, said the validity and clarity of the proposal were straightforward, in that the research would provide needed information on a potentially hazardous substance. He rated the research as moderate-to-high, noting that it would be higher if significant internal human dose could be documented or predicted following oral or other exposure, or if chronic versus episodic exposure could be documented. He agreed that *in vitro* screening batteries such as GeneGo should be employed to further characterize the compound's biological activity. Dr. Rudel asked about the apparently poorly predictive

performance of the QSAR analyses. Dr. Rider responded that GeneGo had predicted some of the positive activities with NBBS that had been confirmed with other methods and confirmed that the training set used with Leadscope was very limited.

Dr. Novak summarized the discussion, saying that the BSC found this project to be of reasonably high priority.

XXVII. NTP Testing Program Research Concept Update: Selected Flame Retardants

A. Presentation

Dr. Mamta Behl updated the BSC on the progress of the testing program for selected flame retardants, specifically the aromatic phosphates (APs) comprising Phase 2 of the research program. She reminded the BSC that these flame retardants had been nominated for NTP testing by the Consumer Product Safety Commission (CPSC) in 2005, and that the BSC had approved the testing in 2006. She said there is extensive exposure to these compounds in the general population through infant sleepwear, upholstered furniture, and plastics, and that some of these flame retardants are being used to replace brominated diphenyl ethers (BDEs), the use of which is being phased out. The CPSC was concerned that it did not have enough toxicity information on these flame retardants to perform regulatory decision-making.

The NTP elected to address the nomination in phases. In Phase 1, which is currently underway, the compounds being studied are antimony trioxide and tris(chloropropyl)phosphate. Phase 3 testing has been deferred; it includes phosphonic acid, (3-((hydroxymethyl)amino)-3-oxopropyl)-,dimethyl ester and tris(hydroxymethyl) phosphine oxide. More information on extractability of those compounds from treated articles is needed prior to making decisions regarding testing needs and study design. Dr. Behl said her update would focus on Phase 2, which includes the APs.

Six compounds in the APs class were nominated because their use is anticipated to increase in upholstered furniture and bedding. In addition to their use as flame retardants, some are used as plasticizers and non-food pesticides. The CPSC requested neurotoxicity, reproductive and developmental toxicity, and subchronic and chronic oral studies.

Dr. Behl reported that the key issues related to these compounds are: (1) commercially available products are often comprised of mixtures containing different structural isomers and other halogenated aryl and aliphatic esters, making it difficult to correlate toxicity with structure; (2) children may be at increased exposure risk due to chewing on treated materials; and (3) chronic toxicity data and developmental and multigenerational toxicity data are lacking.

The NTP research program has two major goals—to study the relative toxicity of the APs as a class, which will include a larger set of compounds, including commercially available mixtures, and comprehensive evaluation of representative compounds with an emphasis on developmental and chronic exposures.

Specific Aim 1 is a class study to evaluate the toxicity of all nominated APs and selected other non-aromatic phosphate esters. This would help determine the relative toxicity and potential

mechanisms of the compounds, as well as characterizing the influence of structural variations. Through Tox21, some HTS data are available to prioritize further *in vitro*, high content, medium-throughput studies of neurotoxicity, steroidogenesis, liver function, and reproductive toxicity. Studies in alternate animal models such as *C. elegans* and zebrafish will also be conducted.

Specific Aim 2 proposes in-depth studies of two specific APs: triphenyl phosphate (TPP) and butylphenyl diphenyl phosphate (BPDP). Those compounds were selected to allow assessment of the comparative toxicity associated with the alkyl-substituted phenyl ring. TPP is extensively present in drinking water and house dust, and has been associated with decreased fertility and alterations in hormones in men. BPDP is also used extensively as a plasticizer in addition to its flame retardant uses, and some studies have shown triggers for reproductive toxicity in animals, although the studies were not extensive or robust. There are some data available on acute toxicity, peripheral neuropathy, and limited subchronic reproductive toxicity. However, chronic studies and studies on developmental toxicity and multigenerational effects are still lacking. Isopropylated phenylphosphate (IPP) may be considered as an alternative or additional compound. Specific Aim 2 will be performed in two phases. Phase 1 will consist of oral toxicity studies in rats and mice. In rats, there will be a dose range-finding study to select doses for a subsequent developmental toxicity study with exposure of dams starting at GD6 and then in the pups. Several toxicity endpoints will be assessed in the offspring. In mice, there will be a subchronic study in adult animals, and ADME/TK studies will also be conducted. Phase 2 will consist of chronic toxicity and carcinogenicity studies in rats and mice.

The studies are expected to: (1) provide information on potential mechanisms and relative toxicity of the APs as a class and (2) provide comprehensive assessment of toxicity and carcinogenicity for TPP and BDPD, to be used by the CPSC for regulatory decision-making.

B. BSC Questions

Dr. Faustman said that she was “intrigued by this class of compounds.” She noted that the effort to substitute another class of flame retardants for those being phased out or banned adds interest to these compounds, particularly since they are organophosphates. She recommended moving the kinetics and QSAR studies “up front,” prior to conducting the other studies, as judging from experience with organophosphates, there may be rich data on kinetics and structure-activity relation that would inform the other studies.

Dr. Birnbaum mentioned that this represented another opportunity for partnering with NIOSH; given the uses of the APs there may be some opportunities for monitoring in the workplace. Dr. Toraason replied that he would look into interest and feasibility within NIOSH.

Dr. Zelikoff asked whether skin hypersensitivity reactions would be assessed, given the high potential for dermal exposure, and wondered why the research was concentrating on oral exposures. Dr. Behl replied that immunotoxicity/hypersensitivity tests are proposed and that the main reason for choosing to concentrate on oral exposures is that children seem to be the most susceptible population, and that is their most common route of exposure. Additionally the design team would revisit the possibility of conducting studies by other exposure routes. Dr. Miller asked about how liver function was to be addressed in the research. Dr. Behl replied that

there might be hepatocyte tests, and potentially livers might be removed from the rats or mice and run through microarray tests. Dr. Birnbaum added that an NIEHS grantee has some data from *in vitro* screens showing that some of the APs being used as flame retardants are potential developmental toxicants.

C. Public Comments

Mr. Manuppello related PETA's comments regarding this research program concept. He said that while PETA commends NTP's proposed use of *in vitro* studies to screen the APs, it questions the decision to conduct in-depth animal testing, particularly on TPP, which has been manufactured for more than 70 years and seems to be data-rich. He cited several studies that had concluded that TPP was of low concern for human toxicity or of low priority for further study.

Regarding BPDP, he urged the BSC to fully consider the relevance of all existing data prior to endorsing potentially duplicative studies that are unlikely to produce new information that would be useful or relevant. He mentioned several previous studies that would be informative on endpoints of interest, including HPV Challenge studies of the commercial form of BPDP. Based upon the existing data for the compounds, Mr. Manuppello said PETA would urge the BSC to reject any proposals for new animal tests for TPP and BPDP.

D. BSC Discussion

Dr. Tracie Bunton, first lead reviewer, said she found the concept to be valid, with a clear rationale. Studies to elucidate the mechanisms of toxicity for this class of compounds would be of considerable value, with the possibility of providing information across programs. She was supportive of Specific Aim 1, but was concerned about the plan for extensive animal testing in Specific Aim 2, especially given the Tox21 initiatives that had been heard in the meeting previously. She felt that there needed to be some justification for going forward with extensive animal testing, particularly given the goals expressed in Specific Aim 1.

Dr. Behl acknowledged the current "juncture" between Tox21 and more traditional animal-based studies. She noted that all of the compounds were nominated by the CPSC, and were seen as equally important. She said with that in mind, rather than conducting in-depth animal studies with all of them, there was sufficient information available regarding potential exposures and toxicity outcomes to select two compounds for comprehensive assessment. Dr. Bunton wondered when Tox21-type studies would be enough. Dr. Birnbaum noted that Tox21 is already being used and would become more valuable for screening and prioritization, but that it would still be a while until the patterns of responses are sufficiently validated to completely obviate the need to perform *in vivo* testing in an animal model. Dr. Bucher added that ultimately it would come down to when the BSC might tell NTP that animal studies are no longer needed because of Tox21. Dr. Miller agreed with Dr. Birnbaum, and advocated more aggressive parallel testing. Dr. Bucher said the 10,000-chemical library would address it.

Ms. Rudel asked Dr. Behl to address the comments made by Mr. Manuppello, pointing out the existence of considerable data on the gaps being addressed by this program. Dr. Behl said some of the previous studies were designed to assess hypersensitivity for computer

components, which is very different from exposures through house dust and children chewing directly on flame retardant treated materials. Also, she said, there were no existing studies of developmental neurotoxicity or multigenerational reproductive effects. The studies are not duplicating any existing projects, and are distinctive. Ms. Rudel felt that it is in NTP's interest to be very clear about recognizing what's out there in terms of existing data, and explicit about what does not exist, and why any particular study is being conducted.

Dr. Novak summarized the discussion, stating his sense was the BSC considers this program a high priority area for future exploration, with a focus on developmental neurotoxicity.

XXVII. Adjournment

Dr. Novak closed the meeting by thanking everyone on the BSC and at NTP, saying his service had been "a fantastic experience."

Dr. Birnbaum thanked the BSC for its hard work over the two "grueling, but extremely exciting" days of the meeting, and thanked her staff for the excellent presentations they had made. Dr. Bucher thanked the BSC for its "usual outstanding job." Dr. White stated that the BSC would be preparing a report on the BSB.

Dr. Novak adjourned the meeting at 4:30 PM.