Department of Health and Human Services
National Institutes of Health
National Institute of Environmental Health Sciences
Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Special Emphasis Panel

Minutes of the Expert Panel Meeting on the Frog Embryo Teratogenesis Assay—*Xenopus* (FETAX): A Proposed Screening Method for Identifying the Developmental Toxicity Potential of Chemicals and Environmental Samples

May 16-18, 2000

The meeting of the Expert Panel on the Frog Embryo Teratogenesis Assay—*Xenopus* (FETAX): A Proposed Screening Method for Identifying the Developmental Toxicity Potential of Chemicals and Environmental Samples was convened on May 16-18, 2000 at 8:00 am, at the Sheraton Imperial Hotel, 4700 Emperor Blvd, Durham, North Carolina 27703, U.S. The meeting was open to the public. Drs. George Daston and Elaine Faustman presided as Co-Chairs. The stated objectives of this Expert Panel Meeting were as follows:

- Develop consensus on the current validation status of FETAX as a screening method for identifying developmental toxicants as described in the FETAX Background Review Document (BRD);
- Provide recommendations for FETAX protocol modifications that might further optimize its use for the evaluation of single compounds, complex mixtures, and environmental samples, and protocol revisions that should be investigated for their potential to enhance the accuracy and reliability of FETAX;
- Evaluate and, as appropriate, recommend revised decision criteria (e.g., a different set point, inclusion of confidence limits, the use of characteristic malformations versus all malformations) that should be investigated further for their potential to enhance the accuracy and reliability of FETAX;
- Discuss the mechanistic relationship (and documented similarities and differences) between the types of malformations induced in *Xenopus laevis* and those induced in laboratory mammals and humans by the same agent;
- Provide recommendations for validation studies (including those incorporating metabolic activation methods, if appropriate) that should be conducted to further evaluate the usefulness and limitations of FETAX as a screening method for teratogens and developmental toxicants;
- Recommend the current and potential uses of FETAX for human health hazard assessment of single compounds, mixtures, and water/soil/sediment samples;
- Review potential uses of other assays utilizing *Xenopus* (e.g., reproductive toxicity assay, tail resorption assay, vitellogenin assay) and
  - Recommend if and how the use of FETAX might be integrated with such methods; and
  - Recommend how further validation studies for FETAX might be linked to these assays, or how these assays might be linked to FETAX;
- Recommend additional research that would support improved understanding of the mechanistic relationship between development in *Xenopus* and in mammals (including

humans), including similarities and differences with regard to response to developmental toxicants; and

- Provide recommendations for further test method development efforts that should be considered to incorporate more mechanistically based assessments, such as alterations in critical gene expression, and that might provide improved test method performance and reliability.

The proposed product of the meeting was an expert panel report evaluating the current validation status and future directions of FETAX, which will be made available to regulatory agencies and other interested parties. The report also will recommend ways to further optimize and validate the FETAX assay.

The meeting was coordinated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and was sponsored by the National Institute of Environmental Health Sciences (NIEHS), the National Toxicology Program (NTP), and the U.S. Environmental Protection Agency (EPA).

Expert Panel Members present included:

- George Daston, Ph.D., Procter & Gamble, Cincinnati, OH (Panel Co-Chair)
- Elaine Faustman, Ph.D., D.A.B.T., University of Washington, Seattle, WA (Panel Co-Chair)
- Gerald Ankley, Ph.D., U.S. EPA, Duluth, MN
- Giovanni Bernardini, Ph.D., Universita dell' Insubria, Varese, Italy
- Chuck Bonham, Ph.D., Colorado State University, Fort Collins, CO
- Michael Brabec, Ph.D., East Michigan University, Ypsilanti, MI
- Nigel Brown, Ph.D., University of London, London, United Kingdom
- Terry Brown, Ph.D., Johns Hopkins University, Baltimore, MD
- Donna Farmer, Ph.D., Monsanto Company, St. Louis, MO
- Anthony Frank, D.V.M., Ph.D., D.A.C.P., D.A.B.V.T., D.A.B.T., Colorado State University, Fort Collins, CO
- Jay Gandy, Ph.D., University of Arkansas for Medical Sciences, Little Rock, AR
- David Gardiner, Ph.D., University of California-Irvine, Irvine, CA
- Robert Grainger, Ph.D., University of Virginia, Charlottesville, VA
- Barbara Hales, Ph.D., McGill University, Montreal, Quebec, Canada
- Bryan Hardin, Ph.D., National Institute for Occupational Safety and Health, Atlanta, GA
- Joseph Haseman, Ph.D., NIEHS, Research Triangle Park, NC
- Robert Hoke, Ph.D., DuPont Health and Environmental Excellence Center, Newark, DE
- Ronald Hood, Ph.D., University of Alabama, Tuscaloosa, AL
- Sidney Hunter, Ph.D., U.S. EPA, Research Triangle Park, NC
- Susan Hurt, Ph.D., D.A.B.T., Rohm and Haas, Spring House, PA
- Carl Keen, Ph.D., University of California-Davis, Davis, CA
- Gary Klinefelter, Ph.D., U.S. EPA, Research Triangle Park, NC

- Sherry Krest, M.S., U.S. Fish & Wildlife Service, Arlington, VA
- Joseph Lary, Ph.D., Centers for Disease Control and Prevention, Atlanta, GA
- Gregory Linder, Ph.D., Oregon State University, Salem, OR
- David Lovell, Ph.D., Pfizer Center Research, Sandwich, England
- Ellen Mihaich, Ph.D., D.A.B.T., Rhodia, Raleigh, NC
- Richard Miller, M.D., University of Rochester, Rochester, NY
- Robert Moore, Ph.D., University of Wisconsin, Madison, WI
- Drew Noden, Ph.D., Cornell University, Ithaca, NY
- Brent Palmer, Ph.D., University of Kentucky, Lexington, KY
- Catherine Price, Ph.D., Research Triangle Institute, Research Triangle Park, NC
- John Rogers, Ph.D., U.S. EPA, Research Triangle Park, NC
- Tom Sabourin, Ph.D., Pro-2-Serve, West Palm Beach, FL
- James Schardein, M.S., WIL Research Laboratories, Ashland, OH
- Jennifer Seed, Ph.D., U.S. EPA, Washington, DC
- Horst Spielmann, M.D., ZEBET, Berlin, Germany
- Takashi Tanimura, M.D., Kinki University, Osaka, Japan

Members of the public present included:

- Andrew Ballard, Bureau of National Affairs, Inc., Washington, DC
- Jim Dumont, Ph.D., Oklahoma State University, Stillwater, OK
- Bruce Ruoft, Janssen, Titusville, NJ
- Eric Wilson, People for the Ethical Treatment of Animals (PETA), Norfolk, VA

Members of ICCVAM, NICEATM, and the ICCVAM Developmental Toxicity Working Group (DTWG), as well as invited speakers included:

<u>ICCVAM</u>

- Angela Auletta, Ph.D., U.S. EPA, Washington, DC
- William Stokes, D.V.M., NIEHS, Research Triangle Park, NC (ICCVAM Co-Chair)

<u>NICEATM</u>

- Sue Brenzel, ILS, Inc., Research Triangle Park, NC
- Loretta Frye, NIEHS, Research Triangle Park, NC
- Tom Goldsworthy, Ph.D., ILS, Inc., Research Triangle Park, NC
- Karen Haneke, M.S., ILS, Inc., Research Triangle Park, NC
- Christina Inhof, M.S.P.H., ILS, Inc., Research Triangle Park, NC
- Linda Litchfield, ILS, Inc., Research Triangle Park, NC
- Barry Margolin, Ph.D., Consultant – Biostatistics, Research Triangle Park, NC
- Debbie McCarley, NIEHS, Research Triangle Park, NC
- Raymond Tice, Ph.D., ILS, Inc., Research Triangle Park, NC

DTWG

- Patricia Bittner, M.S., U.S. Consumer Product Safety Commission, Bethesda, MD
- Jim Burkhart, Ph.D., NIEHS, Research Triangle Park, NC
- Thomas Flynn, U.S. Food and Drug Administration, Laurel, MD
- Gloria Jahnke, Ph.D., NIEHS, Research Triangle Park, NC
- Dennis Lynch, Ph.D., National Institute for Occupational Safety and Health, Cincinnati, OH
- David Morse, U.S. Food and Drug Administration, Rockville, MD
- William van der Schalie, Ph.D. U.S. EPA, Ft. Detrick, MD

Invited Speakers

- John Bantle, Ph.D., Ohio University, Athens, OH
- Doug Fort, Ph.D., Stover/Oklahoma State University, Stillwater, OK
- Joan-Albert Vericat, Ph.D., Sanofi Synthelabo, Gargenville, France

Other Federal employees present included:

- Frank Johnson, Ph.D., NIEHS, Research Triangle Park, NC

OPEN MEETING

Call to Order

Dr. George Daston, Expert Panel Meeting co-chair, called the meeting to order at 8:00 a.m. and asked each person in attendance to state his or her name and affiliation. Dr. Daston informed the participants that the public would be given the opportunity to speak and that anyone addressing the group to please state their name for the benefit of the transcriptionist.

Welcome from the National Toxicology Program

Dr. George Lucier, Director of the NTP, thanked the co-chairs and the invited experts for their time and effort in evaluating the validation status of FETAX and its potential applications in the regulatory arena. Dr. Lucier then briefly discussed the importance of the ICCVAM process in ensuring that alternative tests are appropriately validated in hazard identification and risk assessment. He concluded by emphasizing the significant scientific role NIEHS/NTP has played in the development of alternative tests, including FETAX.

The ICCVAM Test Method Evaluation Process

Dr. William Stokes, ICCVAM Co-Chair, Director of NICEATM, and Executive Secretary for the meeting explained policies and procedures regarding confidentiality and avoidance of conflict of interest situations. Next, he explained the ICCVAM test method review process and the steps that are undertaken in the review of any alternative assay. Dr. Stokes discussed the role of the ICCVAM committee, its expert subgroup (DTWG) and the expert panel, and

the process by which proposed test methods are reviewed and forwarded to agencies for action.

Public Law 103-43 directed the NIEHS to: develop and validate alternative methods that can reduce or eliminate the use of animals in acute or chronic toxicity testing; establish criteria for the validation and regulatory acceptance of alternative testing methods; and recommend a process through which scientifically validated alternative methods can be accepted for regulatory use. Criteria and processes for validation and regulatory acceptance were developed in conjunction with 13 other Federal agencies and programs with broad input from the public. These are described in the document "Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the Ad Hoc Interagency Coordinating Committee on the Validation of Alternative Methods," NIH Publication 97-3981, March 1997. This document is available via the internet at http://ntp-server.niehs.nih.gov/htdocs /ICCVAM.htm.

ICCVAM was subsequently established in a collaborative effort by NIEHS and 13 other Federal regulatory and research agencies and programs. The Committee's functions include the coordination of interagency reviews of toxicological test methods and communication with stakeholders throughout the process of test method development and validation. The following Federal regulatory and research agencies and organizations are participating in this effort:

- Consumer Product Safety Commission
- Department of Defense
- Department of Energy
- Department of Health and Human Services
  - Agency for Toxic Substances and Disease Registry
  - Food and Drug Administration
  - National Institutes of Health
    - National Cancer Institute
    - National Institute of Environmental Health Sciences
    - National Library of Medicine
- National Institute for Occupational Safety and Health/Centers for Disease Control and Prevention Department of the Interior
- Department of Labor
  - Occupational Safety and Health Administration
- Department of Transportation
  - Research and Special Programs Administration
- Environmental Protection Agency

Dr. Stokes then briefly reviewed the timeline for FETAX beginning in 1998 with the request from the U.S EPA for ICCVAM to evaluate the validation status of FETAX. FETAX is a 96-hour assay, which was developed to assess developmental toxicity, and has been used in both human health and ecological assessments. Potential regulatory applications for human health evaluations for developmental toxicity include screening and prioritizing compounds for further testing, evaluating complex mixtures in environmental samples, and providing

supplemental information in a weight-of-evidence evaluation of human developmental toxicity hazards. The U.S. EPA requested that ICCVAM review the validation status of FETAX for various applications, and to determine critical research, development, and validation efforts needed to improve the method. Based on the information available, ICCVAM decided to convene an Expert Panel Meeting. Such meetings are typically convened to evaluate the validation status of a method following the completion of initial development and pre-validation studies. The general objectives of this ICCVAM/NICEATM Expert Panel Meeting were as follows:

- evaluate the current validation status of FETAX,
- recommend research and model development efforts that might improve the performance characteristics (i.e., accuracy, reproducibility) of FETAX for its intended purpose, and
- recommend validation studies needed to further characterize usefulness and limitations and to fill data gaps with regard to chemical/product classes study design and reference chemicals.

The results of the Expert Panel Meeting will be forwarded by ICCVAM to Federal agencies for consideration, and a report of the meeting will be made available to the public.

Regulatory Agency Requirements for Developmental Toxicity Data and Use in Risk Assessment

Dr. Jennifer Seed presented an overview of the U.S. EPA's Guidelines for Developmental Risk Assessment, describing various aspects of developmental toxicity, the pre- and post-natal toxicity associated with the timing of exposure, and considerations for the dose-response relationship. Next, Ms. Patricia Bittner summarized applicable regulations within the U.S. Consumer Product Safety Commission (CPSC) and the impact of the Federal Hazardous Substances Act (FHSA) on testing that might employ FETAX or other alternative tests.

Review of the FETAX Protocol

Dr. Jack Bantle summarized the technical aspects of the FETAX protocol, beginning with a brief review of the history of the assay since its initial development by Dr. Jim Dumont in the mid-1970s through the development of a comprehensive guideline for conducting FETAX published in 1991 under the auspices of the American Society for Testing and Materials (ASTM), as a "Standard Guide for Conducting the Frog Embryo Teratogenesis Assay—*Xenopus* (FETAX)," Annual Book of ASTM Standards, Designation E1439-91, and the publication of a revised ASTM FETAX Guideline (Designation E 1439-98) in 1998. The assay was originally designed as a screening assay for mammalian teratogens, but *X. laevis* has been used for other applications (e.g., endocrine disruptor activity). Dr. Bantle discussed the various individual steps of the assay from breeding to data collection and interpretation. This discussion included a description of the process for the double selection of embryos, the various developmental stages of *X. laevis* during the first 96-hours after fertilization, and the laboratory methods used to expose the developing embryos to test substances with and

without an exogenous metabolic activation system. Subsequently, Dr. Bantle reviewed the methods used for evaluating embryonic mortality and malformations and the various single and multiple criteria used to identify test substances as teratogenic or non-teratogenic in this assay. Dr. Bantle then briefly discussed various laboratory and *in situ* methods for using FETAX to test for the presence of teratogens in soils and sediments. He closed his presentation with examples of data obtained from environmental studies using FETAX.

Summary of Current FETAX Database

Dr. Raymond Tice summarized the current FETAX database developed by NICEATM during preparation of the Background Review Document (BRD). First described were the various single and multiple decision criteria used historically to decide whether a test substance was positive or negative in FETAX. Next, Dr. Tice reviewed the number of studies conducted (276) to evaluate the teratogenic activity of defined substances using FETAX, the number of defined substances tested with (137) and without (35) metabolic activation, the number of defined substances tested in single or multiple laboratories, and the range of data generated by major chemical and product classes. He next summarized the environmental sample database in regard to the number of studies conducted (10), the number of environmental samples tested without metabolic activation only (124), and the lack of replicate studies conducted within and across laboratories. Dr. Tice then summarized the five FETAX validation studies, the methods used to assess reliability (inter-laboratory reproducibility, intra-laboratory repeatability) within each validation study, the extent of concordance across laboratories for each test substance used in a validation study, and the performance characteristics (i.e., accuracy, sensitivity, specificity, false negative rate, false positive rate) of FETAX against laboratory mammalian (rat, mouse, rabbit) and human data, where available. He completed his presentation with a brief review of historical positive and negative control data generated in different laboratories and the results of NTP Quality Assurance Unit data quality audit of the most recent FETAX validation study.

Related *Xenopus* Toxicity Test Methods: Current and Proposed Applications

Dr. Douglas Fort provided a concise review of other current and proposed applications for *X. laevis*. These include a number of other *X. laevis*-based assays under development to identify substances or environmental samples that may disrupt endocrine function (the *Xenopus* Tail Resorption Assay and the Vitellogenin Assay), tests for assessing reproductive toxicity, and tests for exploring limb mal-development, including possible mechanisms of action (*Xenopus* Limb Bud Assay). He also mentioned the use of *X. laevis* in nutritional toxicology studies. Dr. Fort closed his presentation by acknowledging that these developing test methods required appropriate validation, but that they appeared to be useful for the various applications.

Sanofi-Synthelabo Validation Study – Unpublished Data

Dr. Joan-Albert Vericat presented FETAX data generated in his laboratory during an acceptability study of this assay as a screen for mammalian teratogens. The assay was selected because of its relative cost-effectiveness, its ease of application, the relative short-

study duration, and because the developing embryo completed all steps of organogenesis during the exposure period. The utilized protocol differed in several minor ways from that method published by the ASTM. Dr. Vericat compared FETAX results for 34 compounds (many of which were proprietary compounds belonging to different pharmacological classes) with teratogenicity results obtained for the same compounds in laboratory mammals. Based on the results obtained, the accuracy was 79%, the specificity was 92%, and the sensitivity was 71%. Dr. Vericat concluded that these values were acceptable for a screening assay when FETAX was used in the present context. He also discussed briefly the observation that FETAX seemed more predictive of teratogenic results in rabbits than in rats or mice, and that osmolarity values above 150 mM were capable of inducing false positive responses.

Comments on Regulatory Perspectives for *In Vitro* Assays

Dr. Elaine Faustman presented a brief overview of several issues related to the regulatory context for developmental toxicity information. These included:

- There is an essential need for the development of new and improved assays, including *in vitro* ones, for assessing developmental toxicity. The complexity of the developmental toxicity endpoint, however, must be considered during this process.
- The developmental toxicity field is increasingly seeking not just yes/no answers to questions of a chemical's potential to produce developmental toxicity, but is asking for more quantitative dose-response information, information on functional deficits, expansion of the time over which developmental changes are assessed, and more accurate kinetics and exposure assessment methods. The framework for the use of data obtained using *in vitro* assays needs be considered in view of these changing needs.
- If the FETAX assay is proposed for use in screening, then the regulatory and application context for screening needs to be clarified. For example, if the use of FETAX is intended for screening of pharmaceuticals for development of drugs, then it may not be necessary to evaluate the validation status of a screening assay against all aspects of the so-called gold standard (animal bioassay or human epidemiological evaluations), but rather against other screening approaches (e.g., structure-activity relationships, general toxicity information).
- The role of an assay like FETAX in environmental applications such as for prioritization of clean-up/remediation or for determining potential for human or ecosystem risks is very different from the role the same assay may have in pharmaceutical drug screening. The respective roles need to be clearly delineated in order to identify correctly the appropriate criteria by which the assay should be judged.

FETAX is the first developmental toxicity assay to be formally evaluated by ICCVAM. As such, there are issues specific to FETAX and issues generally applicable to the evaluation of the validation status of all future proposed alternative developmental assays. These different, but related, needs should be considered during the Expert Panel Meeting.

Breakout Group Review of FETAX and Related Applications of *Xenopus*

Dr. Elaine Faustman, Expert Panel Meeting co-chair, closed the opening plenary session by reminding the Breakout Groups of the goals and the issues that should be addressed before meeting end. She reminded the participants of the large hurdle that needed to be overcome before any *in vitro* test could be considered validated for regulatory decision-making purposes and that this meeting represented an important step in that process. Dr. Faustman also pointed out the need to identify the appropriate role for assay application for *in vitro* screening versus those applications to replace an existing *in vivo* assay. The potential application of FETAX in environmental hazard identification also needs to be evaluated. After opportunity for public comment, the opening plenary session was adjourned at 12:30 pm.

The Expert Panel reconvened at 1:30 pm, at which time Drs. Daston and Faustman directed the various Breakout Groups to meet individually. Each individual Breakout Group met until 5:30 pm, at which time the meeting was adjourned for the day.

At 8:00 am on May 17, 2000, the Expert Panel met briefly in the plenary session to address any administrative issues; at 9:00 am, the individual Breakout Groups moved to their individual meeting rooms. At 4:30 pm, the Expert Panel reconvened in the plenary session to hear brief progress reports. The meeting was adjourned at 5:30 pm.

On Thursday, May 18, 2000, the individual Breakout Groups reconvened at 8:00 am in their individual meeting rooms. At 10:00 am, the Expert Panel met together in the closing plenary session. Dr. Daston called the session to order and requested that each Breakout Group summarize their conclusions and recommendations.

FETAX Protocol Breakout Group
(A. Levin, co-chair; T. Sabourin, co-chair; D. Farmer, rapporteur; G. Bernardini; A. Frank; G. Klinefelter; G. Linder; J. Haseman)

The Protocol Breakout Group first reviewed the current FETAX protocol, as described in the BRD, and concluded generally that the intended uses of FETAX, its mechanistic basis in terms of human development, its general role in an overall strategy of hazard assessment, the proposed range of materials amenable to test and/or the limits of FETAX according to chemical class or physio-chemical factors were described adequately and that information provided on the protocol was complete and accurate. They concluded also that as embryogenesis appeared to be highly conserved across vertebrates, the use of *Xenopus* as a model organism was reasonable. Several topics that were concluded to warrant additional discussion in the BRD included:

- known similarities and differences in modes and mechanisms of action in FETAX compared to human or laboratory mammals;
- ecotoxicological endpoints as a point of comparison in ecological applications;
- optimization of the metabolic activation system; and
- other possible regulatory uses upon assay validation.

Subsequent to their review of the accuracy of the information provided in the BRD, the Group concluded that there were a number of weaknesses inherent to FETAX. A major limitation of FETAX is that it appears to be sensitive to compounds that alter pH, osmolality, and other properties of the culture water during incubation. In mammals, there may be homeostatic mechanisms that regulate these processes more thoroughly, thus protecting the embryo from these non-specific effects. The Group recommended modifications/improvements to the FETAX protocol in five areas—animal husbandry, training and communications, experimental design, endpoints, and data analysis.

For animal husbandry, the Group concluded that improved control over diet, diseases, housing density, animal source, and the tracking of reproductive history would decrease variability and improve assay reliability. It was noted that *X. tropicalis* appeared to provide some advantages over *X. laevis* on a genetic basis and in terms of average clutch size, time to maturation, and other considerations; thus substitution of this species into the FETAX paradigm should be investigated.

For training and communications, the Group recommended strongly that additional procedures be developed to ensure an adequate level of training for the identification of malformations. This training could include a general workshop among test users conducted on an annual basis (at a minimum) to disseminate information, in-lab training at an expert laboratory prior to performance of this assay, and the Atlas of Abnormalities expanded to include recommended changes in endpoint measurements and subsequently placed on the internet.

In terms of the experimental design, the Group recommended increasing the number of multiple mating pairs per study to allow for a greater assessment of the genetic variability among females. They recommended also that the evaluation of malformations in embryos routinely incorporate a QC Review/Peer review and that an expert panel be convened to retrospectively review the malformation diagnoses in several representative "problem studies" in which there was disagreement between expert pathologists/evaluators scoring the same embryos. The Group recommended that additional research be conducted to evaluate the effect of exposure volumes and embryo loading on the incidence and types of malformations detected in FETAX. Also, if FETAX is were to be used for regulatory purposes, a concurrent positive control, which produces the adverse effects for which the test substance is being evaluated, should be included.

The Group concluded that if the ultimate goal of the FETAX model is to utilize FETAX data to accurately predict the likelihood that a given chemical exposure will produce developmental toxicity in a mammalian species *in vivo*, the endpoints measured in FETAX, as well as the manner in which these endpoints are evaluated, must be expanded and modified to be more compatible with the endpoints of developmental toxicity assays (both *in vitro* and *in vivo*) currently used in regulatory decision making. These data must include the developmental stage of the embryo at the time the assay is completed, embryo length, malformations, functional deficits, and mortality; all data should be recorded at the level of the individual embryo.

With regard to the adequate and meaningful analysis of FETAX data, the Group recommended that all data be tracked and analyzed at the level of the individual embryo. Further, they recommended that the EC50s and LC50s, calculated by the Spearman-Karber or probit methods, be based on a statistically significant increase in malformations (or embryolethality) and on appropriate extrapolations. The Group also suggested the specific method used for assessing the dose-response trend (e.g., Probit or Spearman-Karber) should be stated explicitly in each study and that a formal statistical comparison of slopes for the mortality data and the malformation data be conducted before calculating the teratogen index (TI; i.e., the TI should be calculated only if the slopes are parallel). In addition, the Group recommended that the TI be further evaluated as to whether it is the most appropriate measure of teratogenic response and that the minimum concentration to inhibit growth (MCIG) be further evaluated as to whether it is a useful endpoint.

While not the specific charge of this Breakout Group, they recommended in light of the endocrine disruption arena, the human genome project, and emerging genomic and proteomic technologies, that the framework for a tiered-endpoint battery and a tiered exposure paradigm be conceptualized. If, upon adequate revision, the FETAX model successfully predicts chemicals that are positive in developmental toxicity tests in mammals, a subsequent test might be performed to generate tissues for gene and protein analyses; if the proteins expressed (translated) are also compromised, these genes/proteins may be candidate biomarkers of effect.

FETAX Reliability Breakout Group
(S. Hurt, co-chair; R. Miller, co-chair; R. Hood, rapporteur; E. Mihaich; J. Rogers; H. Spielmann; D. Lovell)

The Reliability Breakout Group concluded that the BRD was well developed, with appropriate detail where possible. Most or all of the difficulties with the BRD reflected both the limitations of the current data sets and the biology available concerning the animal model (e.g., genetic variability, common developmental pathways, and mechanisms for inducing malformations). However, this Breakout Group suggested that an index of important terms (i.e., key words) would aid in finding specific information in the BRD, that it would be helpful if a sample protocol including a flow diagram of the experimental designs and a sample data package were included in the BRD, and that the description of the ASTM statistics – *h* and *k* be expanded with an illustrative diagram.

The Group concluded that FETAX does appear to be capable of measuring key relevant developmental toxicity endpoints, including lethality, malformations, and growth, and of estimating the dose-response relationship for these endpoints. However, based on the available data, the Group concluded that the FETAX results were excessively variable, both within and between laboratories. This excessive variability casts serious doubt on the credibility and usefulness of the assay as a developmental toxicity screen. Variability was observed for all of the major measures reported in the assay, but was particularly marked for the MCIG. Because of this variability, the Group recommended that FETAX not be used for regulatory decision-making at this time. The Group commented that FETAX results

inappropriately focus on teratogenicity, whereas developmental toxicity also includes embryolethality, growth retardation, and functional deficit.

The Group recommended that, after significant protocol modifications are made, the reliability of FETAX should be re-evaluated in blinded tests with coded compounds representing a large number of chemical and mechanistic classes, and from which a range of toxic outcomes would be expected. There should be a scientifically based rationale for the number and nature of compounds selected for testing and for the number of laboratories conducting comparable studies.

FETAX Performance Breakout Group
(N. Brown, co-chair; C. Price, co-chair; B. Hardin, rapporteur; J. Lary; J. Schardein; J. Seed; T. Tanimura)

The Performance Breakout Group stated that they were uncertain about the premise behind the role of FETAX in screening. In particular, the logic behind any decisions that would be made on the basis of FETAX testing was not explicitly stated anywhere in the BRD. In addition, the approach to extrapolation of test chemical concentrations in FETAX to various exposure scenarios in humans or other mammals is lacking and the developmental phases covered by FETAX is too limited to form the basis of regulatory decisions. Thus, the Group stated that FETAX has no current role in regulatory human health risk assessment.

The major conclusion of the Group was that the current data, as comprehensively summarized in the BRD, does not permit any definitive conclusions on the performance characteristics of FETAX. This conclusion was based on the inadequacy of the *in vivo* reference data in the BRD, that the performance comparisons did not take into account the fact that a mammalian prenatal developmental toxicity study is not just a teratogenicity assay, and that there is no objective evidence that the endpoints that are currently being used in FETAX are actually the most predictive of *in vivo* developmental toxicity. Based on this conclusion, the Group recommended that all current performance tables, the associated text, figures, and appendix should be removed from the BRD before publication.

To make use of the existing data, the Group recommended that chemicals for which reliable FETAX mortality, malformation, and growth curves are available should be identified. For these chemicals, acceptable *in vivo* data (excluding human data) should be assembled. From this *in vivo* data, effects on mortality, malformations, and growth should be extracted, and the no observable adverse effect level and lowest observable adverse effect level determined. Next, the FETAX mortality, malformation, and growth data (and the existing TI and MCIG ratios) should be compared with these *in vivo* data. This comparison should be done for each manifestation of developmental toxicity separately and also for all combined. Attempts to correlate specific malformations found in FETAX with those found in other species are not likely to be informative; therefore, such correlations need not be a part of future performance evaluations. These comparisons should be guided by biostatistical expertise. Clearly, these analyses must take into account the fact that *in vivo* embryonic exposure is limited, usually by maternal toxicity, while there is no such exposure limitation in FETAX. The development of a prediction model, from the analyses described above, may help to define an exposure

limit for FETAX. Another possibility may be to incorporate cytotoxicity data from another system, for example cell culture, to relate to FETAX effective concentrations.

The Group suggested that if FETAX were to be considered as an alternative model for all manifestations of mammalian developmental toxicity, the assay would need to be expanded to include assessment of functional deficiencies during later stages of development.

Finally, the Group concluded that if FETAX were to be developed for potential use in regulatory human health risk assessment, further formal validation studies conforming to current ICCVAM guidelines would be required.

Environmental Applications Breakout Group
(R. Hoke, co-chair; R. Moore, co-chair; J. Gandy, rapporteur; S. Hunter; S. Krest; C. Bonham)

The Environmental Applications Breakout Group evaluated the potential for FETAX to serve as a tool for ecotoxicological assessment. They did not address the application of FETAX results from environmental samples to assessments of human health risks. An evaluation of environmental applications of FETAX was difficult because the BRD focused almost exclusively on an evaluation of FETAX relative to mammalian teratogenicity. No comparisons were made of the relative sensitivity of FETAX versus tests with other aquatic test species (i.e., does FETAX provide additional sensitivity, endpoints, exposure routes, or mechanisms not detected by other standardized tests?). The Group concluded that if the use of FETAX in ecological risk assessments is to be evaluated, expansion of the BRD is required to present these analyses.

The Group expressed concern that potential, non-contaminant confounding factors may have been present in FETAX assays on environmental samples. Levels of trace elements, dissolved oxygen, non-trace minerals, redox potential, pH, carbonates, nitrates, ammonia, and other factors are typically evaluated in ecotoxicological testing, but this has not been routinely conducted for FETAX. The buildup of ammonia is a particular concern; ammonia is known to accumulate, as a result of metabolism, to toxic levels in laboratory studies in which solution volumes are low, as they are in FETAX. The accuracy of the data from assays performed in small volumes may be suspect, and this area of the protocol should be changed unless it can be demonstrated that ammonia is not present at adverse concentrations.

In general, the information on sample collection, processing, and storage was considered adequate. The supporting references that document the acceptability of sample storage for up to two weeks before testing should be added to the BRD. As in all environmental sampling programs, additional information on spatial characterization of sites would be useful. Sampling protocols should adhere to accepted sampling and sample handling methods for environmental samples.

The Environmental Applications Breakout Group concluded that:

- Amphibian toxicity assays similar to FETAX may have utility for environmental applications assuming that they reflect exposure routes, effects on additional

endpoints, sensitivity, or mechanisms of action not evaluated by existing aquatic toxicity tests. The utility of amphibian toxicity assays will be increased if the purpose of the tests can be clearly identified (e.g., ecotoxicity screening, surrogate for other amphibian species, etc.).

- The utility of the existing FETAX protocol relative to existing aquatic toxicity assays can only be determined after compilation and evaluation of existing data from standard aquatic toxicity tests.
- Difficulties exist with specific procedures used in the current test method (e.g., organism biomass loading, test solution volume, lack of standard chemical measurements, exposure regime for hydrophobic or unstable toxicants etc.). These are major issues that may require protocol changes and/or additional research.
- The current use of FETAX in environmental assessments is not adequately consistent with standard aquatic toxicity testing protocols.

The recommendations of this Group included:

- Available FETAX results should be compared with results from standard aquatic toxicity tests. At a minimum, this comparison should include results from acute studies using an invertebrate (*Daphnia magna*, *Ceriodaphnia dubia*) and fish (fathead minnow, rainbow trout, and bluegill), and short-term chronic and chronic studies using an invertebrate (7-day *C. dubia*, 21-day *D. magna*) and fish (7-day fathead minnow, 28-day fathead minnow, 90-day rainbow trout). Comparisons should also be made to all available results from other amphibian toxicity tests.
- The FETAX protocol must be modified as necessary to address methodological questions and concerns relative to organism loading, test solution volume, chemical measurements (water quality and test substances), and exposure regimes for hydrophobic or unstable toxicants.
- Environmental application of the FETAX assay should incorporate standard protocol design considerations found in existing aquatic toxicity tests, such as measurement of routine water chemistry parameters.
- In analysis of data from the FETAX assay, the use of the TI value should be abandoned and individual endpoint measures used for mortality, malformations, and growth.
- The utility of alternative amphibian species (e.g., native species and /or *X. tropicalis*) should be evaluated given that they may offer testing advantages such as shorter generation time, increased fecundity, etc., relative to *X. laevis*.
- Alternative amphibian assays that incorporate new endpoints (e.g., limb bud development, tail absorption) should receive additional research effort.

Research and Development Breakout Group
(T. Brown, co-chair; B. Hales, co-chair; M. Brabec, rapporteur; D. Gardiner; R. Grainger; G. Ankley; C. Keen; D. Noden; B. Palmer)

The Research and Development Breakout Group concluded that a critical assessment of the organism is necessary for FETAX. This assessment involves an evaluation and comparison of *Xenopus* species, the presently used tetraploid, *laevis*, as well as the diploid, *tropicalis*. The potential advantages of *X. tropicalis* included:

- a diploid genome;
- can be used in mutagenesis and knockout technologies;
- gene dosage effects that may increase sensitivity and reproducibility (gynogenetic *laevis* may provide an opportunity for parallel analyses);
- higher normal temperature range during development leads to more rapid embryonic development (2 +/- days vs. 4 days for *laevis*), more rapid generation of reproductive adults (3-4 months for *tropicalis* vs. 1+ yr. for *laevis*), and is closer to the functional range for the microsomal activation system;
- embryos are smaller, thereby allowing more embryos per dish;
- readily accessible with some 5^{th} generation inbred animals now available;
- some members of the *Xenopus* developmental biology research community are now using *tropicalis* and data to date indicate that the extensive mechanistic database available for *laevis* is applicable to *tropicalis*; and
- the suitability of *Xenopus* for transgenic studies provides an opportunity to incorporate a range of reporter constructs to meet the unique and specific needs of each context (e.g. signal pathways, regulatory genes, tissue/organ specific genes, and metabolic pathway genes).

The Group further recommended that:

- whichever species is used in FETAX, there must be a consistent, standardized, reliable source and supply of animals;
- animal housing and diet should be standardized;
- temperature should be controlled to maintain rates of development for optimal sensitivity, to improve efficacy of the metabolic activation system, and to avoid extremes that magnify variability;
- the dejellying process should be investigated regarding potential alternatives that will be effective, yet both easier and safer (less “toxic” to embryos) to use;
- the biomass and number of embryos in the test system should be standardized and the relationship of mass per unit volume needs to be optimized;
- the microsomal activation system needs to be assessed for activity throughout the test period and alternative sources (e.g. human cells in co-culture) or strategies (transgenic) should be investigated. Extreme variability in activity may be inherent in the current protocols; and
- technical training for assay procedures and assessment of malformations and other endpoints in screening must be rigorous, intensive and standardized.

The Group concluded that FETAX augmented with application-specific modifications could provide a good system in which the needs of a variety of users can be met. They

recommended that the basic FETAX approach moves beyond positive/negative screening analyses and includes some or many of the following:

- Dose-response relationships must be considered for each chemical. Characterization of malformations must be standardized and considered in detail regarding their 1) time of appearance (especially including gastrula and neurula stages); and 2) specific description according to standardized criteria. Characterization should, whenever possible, identify developmental defects using criteria and categories applied in mammalian systems. Not every application will require detailed assessment of all malformations, but without this information the ability to validate the assay and to decide the best ways of "simplifying" the initial analyses is lost; 3) size data should be digitized according to standard protocols for capturing images; 4) the atlas of malformations should be expanded to include, and, as best as possible, describe all the developmental disruptions that occur spontaneously or experimentally in *Xenopus*; 5) subsets of malformations should be clearly identified that appear similar to, as well as different, from those found in mammals; and 6) the stringency of morphological analyses should be improved.
- An accessible database of FETAX results and malformations is essential. This database should be: 1) web-based; 2) utilize standardized imaging and digitizing protocols; 3) employ standard nomenclature and categorization; and 4) be linked to XEN-base and other pertinent developmental, toxicological, and *Xenopus* internet sites. Development and integration of the preceding suggestions may require assembly of a small, interdisciplinary expert panel to establish guidelines for implementation.

The Group considered several applications of *Xenopus* embryos that could provide important augmentation to existing assays. These included:

- Transgenic methods for *Xenopus* currently available to introduce reporter-linked transgenes. Many of these could greatly enhance both the specificity and ease of assay in sentinel applications.
- Other critical windows of development for which *Xenopus* has shown promise for developmental toxicology applications include limb development, tail resorption, juvenile/adolescent responses, reproductive toxicology including male and female gonadogenesis, germ cell maturation and reproductive behavior, and vitellogenin synthesis.
- FETAX might be expanded in the future to include additional endpoints, such as assays for swimming, light sensitivity and reflexes, somatosensory reflexes, and feeding behavior.
- Gene and gene product expression profiles can be developed; 1) using RNA for cDNA arrays of developmentally critical genes, comparisons between *Xenopus* and mammalian embryos should be made, defining stage, tissue and insult-specific responses; 2) transgenics with reporter genes; 3) *in situ* analyses which draw upon an extensive library of available probes; and 4) proteomics based upon emerging technologies. Data collected from the NIEHS *Xenopus* cDNA microarrays will be useful in this respect.

The use of cDNA microarrays in the context of FETAX is not addressed in the BRD, other than noting that the application of this technology is possible. However, NIEHS has established a resource center that is presently sequencing embryonic stem cells (ESTs) obtained from a normalized library of cDNAs from unfertilized *Xenopus* eggs and similar or related projects using ESTs or cDNA libraries are ongoing at other institutions. These efforts are, by their nature, related to an overall genome project and database being coordinated by various investigators under the auspices of the National Institutes of Health. This will require establishing resource centers to create, distribute, and analyze such microarrays as an experimental tool to be utilized by individual investigators. Although initial efforts will build on the wealth of knowledge to be ascertained in relation to *X. laevis* development, parallel efforts are evolving that will also include *X. tropicalis*. However, it should be noted that the application of cDNA microarray analyses to further improve FETAX may be problematic in light of the variability in FETAX that is apparent in toxic chemical detection and variations that exist among laboratories. At present, the use of cDNA microarray analyses in the field of toxicogenomics is not hypothesis driven and such studies will require establishing appropriate and necessary controls with adequate numbers of replicates to permit comparisons with analyses involving toxic exposure. Developmental stage-specific cDNA libraries with coincident cDNA microarrays over the stage 8-46 window will have to be developed if relevant developmental pathways that evolve during FETAX are to complement the results of toxic exposure on *Xenopus* embryo development. To validate approaches with *Xenopus*, there will be a need to assess gene expression profiles in mammalian embryos in parallel for comparative and correlative purposes.

The Group concluded that there is an excellent potential to develop *Xenopus* as a model organism for developmental toxicology and the probability of such occurring may be directly related to the willingness of various agencies to fund the necessary research that is being recommended.

Public Comments

During the closing plenary session, Dr. Daston asked for public comments. Mr. Eric Wilson, representing PETA, thanked the experts at the meeting, on behalf of PETA, for their hard work and insights. Mr. Wilson commented that there seemed to be a fair amount of uncertainty regarding the future of FETAX and expressed a concern that the standards for the method may be unrealistically high, with too many hurdles that must be overcome before the assay can be considered validated. He pointed out that although the assay may not be useful in all situations and for all chemicals, it might still be useful in some situations or for some chemicals, which would result in reducing animal use. Mr. Wilson stated that the most important lesson from this meeting was the need for much greater financial support from the Federal government for actual test method development and full validation studies. The regulatory agencies, such as the U.S. EPA and the Food and Drug Administration, which require toxicity tests for developmental endpoints have to step forward and provide funding. Also, as assay cost was an issue raised during the meeting, PETA feels strongly that financial consideration should not factor into scientific discussions of alternative test method development. Rather, the decision to use an alternative method instead of an *in vivo* method must be driven by the professed desire of both the government and industry to reduce the

pain and suffering of animals used in laboratory tests. It is not a decision that should be made by the finance department, but by people who realize that we cannot continue using old *in vivo* tests when more humane ways exist. Dr. Daston thanked Mr. Wilson for his comments.

Expert Panel Meeting Conclusions

During the closing plenary session, Dr. Faustman presented several general conclusions as follows:

- The number of expert scientists and the range of disciplines involved in this Expert Panel Meeting on FETAX confirm the interest in and the need for alternative approaches for looking at developmental toxicity.
- At this meeting, it was not only important to develop recommendations that were specific for FETAX, but also to develop recommendations that were more general for moving the field of *in vitro* developmental toxicity assessment forward.
- The meeting demonstrated the tremendous efforts by the FETAX organizers in their attempts to develop, standardize, and validate their protocol. There was general interest among the scientific community in *Xenopus* as an interesting vertebrate model organism for studying developmental pathways and processes.
- The initial assessments of FETAX clearly identified problems in variability and because of this variability, the Panel members have identified additional assessments that are needed across chemicals and laboratories. Panel members developed specific recommendations for addressing these problems by making suggestions for changes in experimental protocol and performance outcomes.
- The Panel members concluded that as currently constituted as a teratogenesis assay, FETAX is not sufficiently validated or optimized to be used for regulatory applications.
- The Panel members have concluded that the number of endpoints considered in FETAX should be expanded to increase an understanding of how FETAX performance might be improved or more fully validated in identifying developmental toxicants.
- The Panel members identified the need to further develop specific application, validation, and utilization criteria for FETAX, and especially to evaluate decision criteria other than the teratogen index (TI). The Panel members reiterated that the TI was not an appropriate metric for evaluating developmental toxicity..
- One critical need identified for any further development and validation of *in vitro* developmental toxicity assays is the development of reliable and accurate comparative databases for animal and human developmental toxicants.
- Another critical need is the development of validation criteria specific for developmental toxicity.

Drs. Faustman and Daston thanked the meeting participants for their involvement in this very important process and ICCVAM and NICEATM for supporting, organizing, and managing this important meeting. Both felt that the meeting resulted in a series of constructive recommendations for moving forward both with a final assessment of FETAX as well as for other alternative developmental toxicity assays.

Dr. Stokes then presented the proposed timeline for completion of FETAX-related activities, including final reports and publication. There was general discussion as to points of contact, logistics, and public involvement in this process.

Dr. Stokes closed the FETAX Expert Panel Meeting by thanking the invited expert scientists for agreeing to participate and for their contributions during the meeting and by thanking the invited speakers and FETAX developers for their contributions. He expressed appreciation to Dr. Angela Auletta for chairing the DTWG during this process and the members of the DTWG for their participation and contributions.

The FETAX Expert Panel Meeting was adjourned at 12:16 pm on May 18, 2000.