Can Alternative Methods Contribute to the Identification of Teratogenic Exposures?

Julieta M. Panzica-Kelly and Karen Augustine-Rauch, *Discovery Toxicology, Bristol-Myers Squibb Lawrenceville, New Jersey*

Introduction

A convergence of *in vitro* approaches, high throughput screening methodology, computational approaches and characterization of toxicity pathways has enabled a revolution in the toxicology field. The National Research Council (NRC) recognized these advancements and published a groundbreaking report "Toxicology in the 21st Century," which envisions how the field may evolve by applying these advancements. In lieu of traditional toxicological assessment evaluating adverse findings in animal studies, the future approach is based upon interpretation of toxicological mechanisms, primarily by using a suite of *in vitro* models for testing compounds and a systems biology approach for interpreting data for signatures indicative of toxicity. This approach is envisioned to accomplish both hazard and risk assessment. Hazard assessment involves the identification of a compound's potential to cause toxicity. In the NRC paradigm, this would be accomplished by assessing the compound's response through an *in vitro* test battery and assessment of its associated profiles of cytotoxicity, cellular changes, and/or aberrations of toxicity pathways. Risk assessment is a process that estimates the nature and probability of adverse effects *in vivo* based upon the degree of exposure to the agent. In the NRC paradigm, risk assessment would be accomplished by assessing the concentration response of a chemical in the *in vitro* test battery and using extrapolation modeling to project animal/human adverse toxicological exposures.

Bringing the NRC vision to fruition is particularly relevant to the developmental and reproductive toxicology (DART) field, where traditional whole animal assessments are estimated to represent about 70% of the animal use and 90% of the costs associated with industrial toxicology assessments. To this end, a number of *in vitro* developmental model systems have been evaluated for refining, reducing, or replacing animals for assessment of teratogenic potential of pharmaceutical compounds and chemicals. Such models have included rodent embryonic stem (ES) cells or human induced pluripotent stem (iPS) cells, which have the potential to be integrated into high throughput screening (HTS). Additional animal-based assays, such as rodent micromass, which models early chondrocyte or neuron development, and rodent whole embryo culture have been under evaluation. Assays based on the use of alternative species, such zebrafish embryo culture, are also of extensive interest because of the ability to test a whole embryo.

These assays have shown promise as hazard identification assays for assessing the teratogenic potential of chemicals. However, a less established area is the ability to use these assays in risk assessment applications, which will require extensive *in vitro* assessment in context of *in vivo* exposure information and eventual generation of pharmacokinetic models that will link *in vitro* concentrations to those expected *in vivo*. The DASTON validation list was published as a means to provide researchers with a test set of compounds that have been tested in standard DART

assessments and include *in vivo* exposure information associated with non-teratogenic and teratogenic exposures. To this end, early efforts have begun to examine developmental toxicology assays for concentration-related response signals of teratogenicity and to compare them against test animal *in vivo* exposures that were not developmentally toxic versus those that produced malformations or were embryolethal. We evaluated the DASTON validation compounds and a set of pharmaceutical compounds to examine the potential relationship of a respective compound's concentration response in mouse ES cells versus the rat whole embryo culture (WEC) assay. Our intent is to tier and/or integrate these assays as a means to evaluate *in vitro* to *in vivo* (IVIV) exposure correlations. The attributes of the ES assay include its potential for HTS, ability to test wide concentration ranges of a compound and its adaptability to be applied to a human iPS model. The rat WEC assay represents an intact whole embryo model representing the primary *in vivo* test species used in DART assessment. In addition, the WEC media is comprised of 70% serum, which is similar to *in vivo* circulation and may better model a respective compound's protein binding and activity profiles. A brief description of the ES and WEC models and an example of how this evaluation may be conducted is presented for the human teratogen valproate.

Mouse Embryonic Stem Cells

Mouse ES cells have been used extensively to classify the teratogenic liability of compounds undergoing early assessments. ES assays typically use the cytotoxicity of fibroblasts or ES cells as well as differentiation endpoints. ES cells are maintained as pluripotent in the presence of media supplements until the start of the assay (Figure 1A). Upon removal of these supplements, the cells aggregate into clusters called embryoid bodies (EBs) (Figure 1B) and spontaneously differentiate into cardiomyocytes. The ES assays typically compare a compound's fibroblast cytotoxicity profile with its effect on inhibiting EB cardiomyocyte differentiation. ES cells are treated with a concentration range of compound and the concentration that causes 50% cytotoxicity (the IC50) is calculated from the cell viability curve (Figure 2). The IC50 is used in combination with differentiation endpoints (beating cardiomyocytes, transcriptional targets of differentiation, etc.) to predict teratogenic liability. Variations of ES assays with diverse differentiation endpoints were developed over the last decade and achieve about 70-80% accuracy in correctly classifying compounds that were not developmentally toxic versus those that produced malformations or were embryolethal. Efforts are in progress to determine whether ES cytotoxicity profiles can predict a range of possible *in vivo* teratogenic exposure concentrations. Working with a set of compounds with characterized *in vivo* teratogenic maximal concentration (Cmax) exposures, a range algorithm was generated for testing larger numbers of compounds. For example, working with ~20 DASTON and pharmaceutical compounds, a preliminary teratogenic Cmax concentration range algorithm was developed using ES IC50 and 1:25 IC50 values as a possible adverse effect exposure range. This algorithm can be applied to predict the exposure range used in DART testing. As an example, the teratogenic Cmax concentration of rats treated with valproate was 800 µM. By using the ES IC50 value (3,643 µM), a range of possible effect concentrations was identified (as 146 - 3,643 µM, Figure 2). The predicted *in vivo* adverse effect range may also be used to identify a concentration range to test in the rat WEC assay to refine prediction of the adverse effect Cmax range.

Figure 1. A. Pluripotent mouse embryonic stem D3 cells. B. Day 4 embryoid bodies.
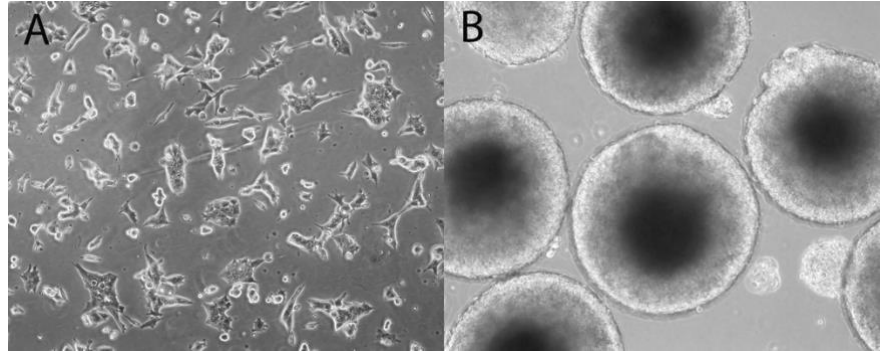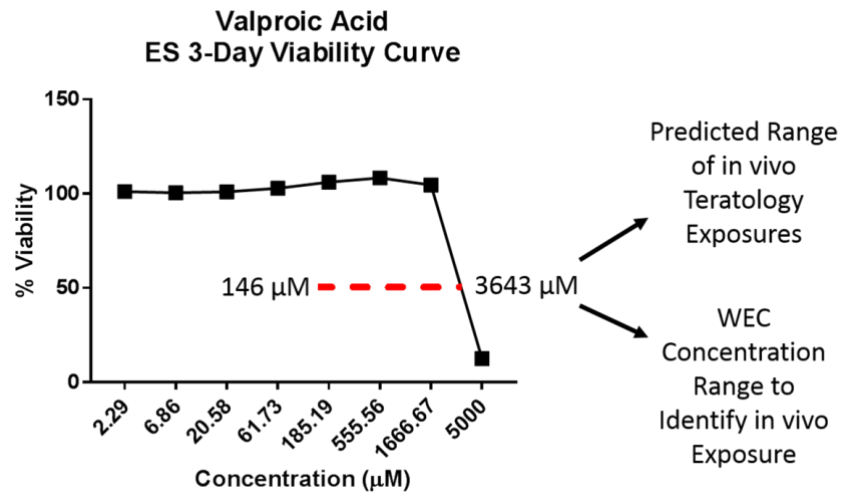

Figure 2. Valproate 3-day ES D3 viability curve. The range of possible *in vivo* teratology exposures may be predicted by calculating the IC50 from this curve (range: 1:25 of the IC50 - IC50). The published *in vivo* effect concentration (800 μM) was captured correctly.

Rat Whole Embryo Culture

Neurulating rat embryos are placed into culture in the presence of compound for 2 days (Figure 3A). During culture, the embryos undergo early organogenesis, where primitive structures and organ systems form (Figure 3B). At the end of culture, embryos are morphologically assessed for alterations in viability, growth, and malformations. These basic endpoints have been applied in various forms of WEC assays for teratogenic assessment, and the approach may also be applicable for identifying potential teratogenic exposures. In most cases the teratogenic Cmax of the ~20 DASTON and pharmaceutical compounds caused a > 30% incidence of malformed embryos, whereas the known negative *in vivo* concentration were comparable to vehicle (<10%). For example, when treating rat embryos with valproate at the teratogenic Cmax concentration (800 μM), 100% of the embryos were severely malformed (Figure 4) but when embryos were treated with 1 μM valproate, which is well below the rat lowest-adverse-effect exposure, only 10% of the embryos were malformed.

Figure 3. Rat embryos. (A) Gestation day 9 rat embryo. (B) Gestation day 11 rat embryo grown *in vitro*.
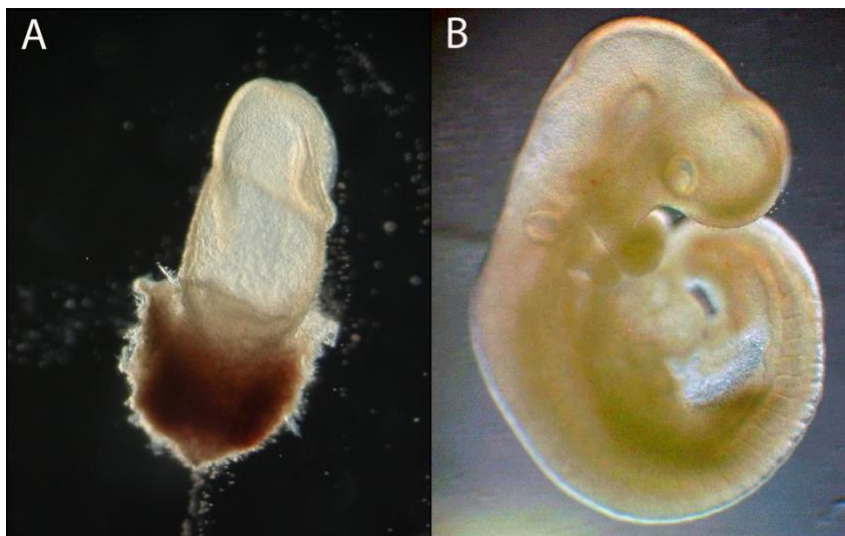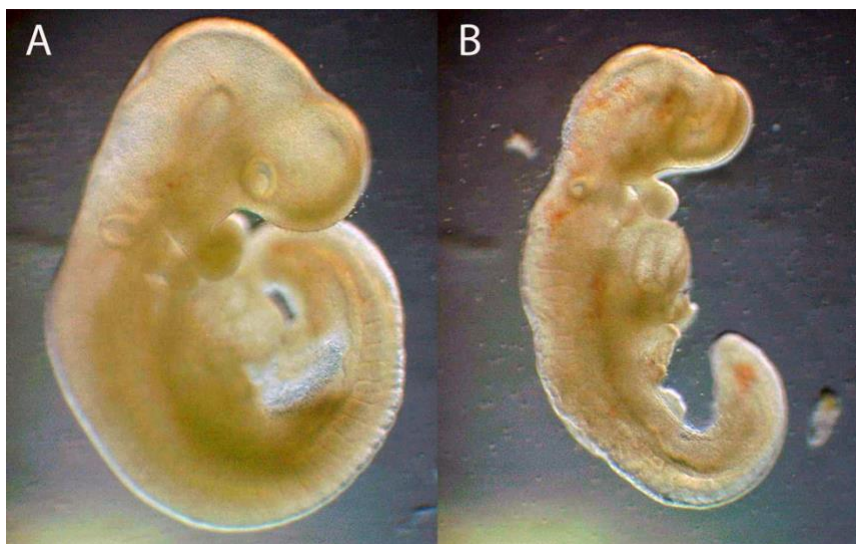

Figure 4. Whole embryo culture gestation day 11 rat embryos cultured in presence/absence of valproate. (A) Vehicle (0.02% DMF) treated embryo with normal morphology. (B) 800 µM valproate treated rat embryo with abnormal morphology (dysmorphology of the spinal cord and forebrain).

It may be possible to integrate the ES IC50 and WEC approach for predicting teratogenic Cmax exposures of compounds that have not been assessed in DART studies. For instance, the ES cytotoxicity profile may be used to identify a teratogenic range that could be further tested in a concentration range in WEC. The adverse effect exposure concentration would be predicted to be greater or equal to the calculated 30% effect concentration.

Concluding Statement

Efforts in developing approaches to improve *in vitro* to *in vivo* exposure relationships is a work in progress. The described model is preliminary in nature but is intended to provide an example of how this effort can be an approach towards achieving the Tox 21 goal of establishing methods to better correlate *in vitro* to *in vivo* adverse effect exposures. Available exposure data of definitive teratogenic exposures is limited in the literature but unpublished resources are available within industrial sectors. Therefore collaboration among scientists in sharing such data and joint efforts to generate and

validate promising methods are key for successfully generating robust models for assessing *in vitro/in vivo* exposure correlations.

Suggested Reading

Daston G, Beyer B, Carney E, Chapin R, Friedman J, Piersma A, Rogers J, Scialli A. Exposure-based validation list for developmental toxicity screening assays. Birth Defects Research (Part B) 2014:00:1-6.

Daston GP, Chapin RE, Scialli AR, Piersma AH, Carney EW, Rogers JM, Friedman JM. A different approach to validating screening assays for developmental toxicity. Birth Defects Research (Part B) 2010: 89: 526-530.

DeJongh H., Forsby A., Houston J.B., Beckman M., Combes R. and B.J. Blaauboer 1999. An integrated approach to the prediction of systemic toxcity using computer-based biokinetic models and biological test methods: Overview of a prevalidation study based on the ECITTS. Toxicology *in vitro* 1999: 3: 549-554.

Hartung T. and C. Rovida C. 2009 Chemical regulators have overreached. Nature 2009: 460(27): 1080-1081.

National Research Council (NRC). Toxicity Testing in the 21st Century: A Vision and Strategy. Washington D.C. 2009: National Academy Press.

Piersma A, Janer G, Wolterink G, Bessems J, Hakkert B, Slob W. Quantitative extrapolation of *in vitro* whole embryo culture embryotoxicity data to developmental toxicity *in vivo* using the benchmark dose approach. Toxicological Sciences 2008:101(1):91-100.

Spielmann H, Pohl I, Doring B, Liebsch M, Moldenhauer F. The embryonic stem cell test (EST), an *in vitro* embryotoxicity test using two permanent mouse cell lines: 3T3 fibroblasts and embryonic stem cells. *in vitro* Toxicology 1997:10:119-127.

Verwei M, van Burgsteden J, Krul C, van de Sandt J, Freidig A. Prediction of *in vivo* embryotoxic effect levels with a combination of *in vitro* studies and PBPK modelling. Toxicology Letters 2006:165:79-87.

Vorhees CL. Teratogenicity and developmental toxicity of valproic acid in rats. Teratology 1987: 35: 195-202.

Zhang CX, Ball J, Panzica-Kelly JM, Augustine-Rauch K. *in vitro* Developmental Toxicology Screens: A Report on the Progress of the Methodology and Future Applications and Future Applications. Chemical Research in Toxicology 2016: