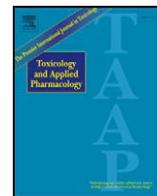




Contents lists available at ScienceDirect

Toxicology and Applied Pharmacology

journal homepage: www.elsevier.com/locate/ytaap

Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics

Paul R. West^{a,*}, April M. Weir^a, Alan M. Smith^a, Elizabeth L.R. Donley^a, Gabriela G. Cezar^{a,b}^a Stemina Biomarker Discovery, Inc., 504 S. Rosa Rd., Suite 150, Madison, WI 53719, USA^b University of Wisconsin-Madison, Department of Animal Sciences, 1675 Observatory Drive, Madison WI 53706, USA

ARTICLE INFO

Article history:

Received 2 February 2010

Revised 11 May 2010

Accepted 13 May 2010

Available online xxx

Keywords:

Teratogen

Developmental toxicity

Stem cells

Biomarkers

Metabolomics

ABSTRACT

Teratogens, substances that may cause fetal abnormalities during development, are responsible for a significant number of birth defects. Animal models used to predict teratogenicity often do not faithfully correlate to human response. Here, we seek to develop a more predictive developmental toxicity model based on an *in vitro* method that utilizes both human embryonic stem (hES) cells and metabolomics to discover biomarkers of developmental toxicity. We developed a method where hES cells were dosed with several drugs of known teratogenicity then LC–MS analysis was performed to measure changes in abundance levels of small molecules in response to drug dosing. Statistical analysis was employed to select for specific mass features that can provide a prediction of the developmental toxicity of a substance. These molecules can serve as biomarkers of developmental toxicity, leading to better prediction of teratogenicity. In particular, our work shows a correlation between teratogenicity and changes of greater than 10% in the ratio of arginine to asymmetric dimethylarginine levels. In addition, this study resulted in the establishment of a predictive model based on the most informative mass features. This model was subsequently tested for its predictive accuracy in two blinded studies using eight drugs of known teratogenicity, where it correctly predicted the teratogenicity for seven of the eight drugs. Thus, our initial data shows that this platform is a robust alternative to animal and other *in vitro* models for the prediction of the developmental toxicity of chemicals that may also provide invaluable information about the underlying biochemical pathways.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Birth defects are a leading cause of postnatal deaths and pediatric disorders. Approximately 3% to 5% of children in the United States exhibit birth defects and of these, between 2% and 3% can be classified as teratogen-induced malformations (Finnell, 1999). Teratogens include a wide range of chemical substances for which currently available animal models have a limited ability to correlate toxicity outcomes to human response (Shanks et al., 2009). In addition, there is a significant global push to reduce the use of animals in toxicity testing. The response of animals to teratogens differs significantly from those of humans; the most striking illustration of this problem occurred when thalidomide was released into the market in 1954 (Miller and Stromland, 1999) for, among others, the treatment of morning sickness in pregnant women. Thalidomide had been tested in rodents, and declared safe for pregnant women; however, an alarming increase in limb defects among newborns was linked exclusively to its use and by 1962, the drug had been removed from the market (Toms, 1962; Brent and Holmes, 1988), but not before the drug had caused more than 10,000 cases of birth defects worldwide.

Thus, there is an unmet need to develop more accurate methods for human developmental toxicity screening.

HES cells were first derived from the inner cell mass of blastocysts (Thomson et al., 1998). Given the human embryonic origin of these cells, we propose that an *in vitro* teratogenicity test using hES cells may produce more accurate human endpoints than animal tests, while at the same time reducing cost and time and increasing predictability. Neurodevelopment is recapitulated and faithfully reproduced by the differentiation of hES cells (Zhang et al., 2001; Keirstead, 2005), thus the detection of metabolites in hES cells should provide insight into mechanisms of injury in early human development. The use of non-human embryonic stem cells to predict developmental toxicity of chemical substances has been previously established. The European Centre for the Validation of Alternative Methods (ECVAM) has approved the mouse embryonic stem cell test (EST) as a teratogenicity screen. This test accurately predicted the teratogenicity of 78% of 20 compounds in a test panel (Genschow et al., 2002). However, the biological mechanisms of the EST are still poorly understood, the test is not specific to human response and the endpoint (number of beating heart cells differentiated from embryoid bodies) is somewhat subjective (Adler et al., 2008).

Metabolomics assesses functional changes in biochemical pathways by detecting changes to the dynamic set of small molecules that

* Corresponding author.

E-mail address: pwest@stemina.com (P.R. West).

comprise the metabolome and has identified physiologically relevant biomarkers for Parkinson's disease, (Bogdanov et al., 2008), early detection of cancer (Kind et al., 2007), resistance to therapy (Serkova et al., 2007) and identified the potential role of sarcosine in aggressiveness of prostate cancer (Sreekumar et al., 2009). These studies have shown that metabolomics delivers sensitive and quantitative biomarkers, while also elucidating biological mechanisms, a key aspect of the hES cell based developmental toxicity assay described herein.

The feasibility of metabolomics in biomarker discovery has been demonstrated by multiple studies (Tan et al., 1998; Sabatine et al., 2005; Cezar et al., 2007). The novel preliminary study by Cezar was the first to propose metabolomics and hES cells for developmental toxicity testing. This platform was able to detect differences in the metabolic profiles of hES cells in response to valproate, revealing novel biochemical pathways of injury that were in agreement with data from previous *in vivo* valproate studies (Vriend and Alexiuk, 1996). Our application of metabolomics to analyze metabolites secreted by hES cells provides good preliminary data of an effective *in vitro* human model to predict developmental toxicity. Extracellular small molecules were analyzed because one of our future goals is to examine biomarkers that we have found *in vitro* in biofluids *in vivo*. Biofluids such as serum, urine and cerebrospinal fluid contain a mixture of extracellular biomolecules and their analysis can be performed using procedures that are less invasive than tissue biopsies.

In the present study, WA09 hES cells were dosed with a training set consisting of 18 drugs of known teratogenicity. The resulting changes in abundance levels of small molecules in the spent media from the dosed cells versus the undosed cells (control) were determined by LC–MS, a standard analytical approach in metabolomics (Goodacre et al., 2004) providing robust identification of small molecules in complex mixtures.

A random forest statistical model was tested in two blinded studies consisting of 8 compounds and correctly predicted the teratogenicity for 7 out of the 8 drugs. Thus, we hereby report on our results for what appears to be the establishment of a highly predictive *in vitro* system for predicting chemical toxicity during early human development. We anticipate that biomarkers of teratogenicity discovered with this method will eventually enable us to develop a more high-throughput, targeted developmental toxicity screen that offers quantitative human endpoints.

Methods

hES cell culture

Maintenance and passaging. WA09 hES cells, obtained from WiCell Research Institute (NIH National Stem Cell Bank, Madison, WI) were cultured in 6-well plates on Matrigel (BD Biosciences, San Jose, CA), in mTeSR1 media (Stem Cell Technologies, Vancouver, BC) incubated at 37 °C under 5% CO₂ in a Thermo Electron Forma Series II Water Jacket CO₂ Incubator. HES cells were passaged every 3 or 4 days at a 1:3 or 1:6 seeding density for routine culture conditions. For dosing experiments, hES cells were plated, allowed to attach and grow for 3 days, then dosed for 4 days, a precedent established in a previous study (Cezar et al., 2007). In order to ensure that hES cells would not require passaging during these 7 days, hES cells were passaged at a low density of 1:10 or 1:12. To passage hES cells, the StemPro® EZPassage™ disposable stem cell passaging tool (Invitrogen, Carlsbad, CA) was used to detach the cells from the wells. Detached cells were removed with a pipette and distributed to new Matrigel plates.

hES cell dosing. A training set of established teratogens and non-teratogens (Table 1) was used to dose hES cells. All tested chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Cells were dosed

Table 1

Chemical compounds in the training and test set (blinds) used for dosing, their classification according to teratogenicity and prediction model incorporation.

Stemina classification	Compound	Model training set	ECVAM classification	FDA classification	
Non-teratogens	Ascorbic acid (AA)	TS1,2	Non-teratogens	A	
	Isoniazid (Is)	TS1,2		C	
	Penicillin G (PG)	TS1,2		B	
	Folic acid (FA)	TS1,2	Not classified	A	
	Levothyroxine (Ly)	TS1,2		A	
	Retinol (Re) (Blind 1)	TS2		A	
	Doxylamine (Do) (Blind 2)	TS2		A	
	Thiamine (Th) (Blind 8)	TS2		A	
	Aspirin (As)	TS2		Weak/moderate teratogens	C
	Caffeine (Caf)	TS2			B
Diphenhydramine (Dh)	TS2	B			
Dexamethasone (Dex)	TS2	C			
Teratogens	Diphenylhydantoin (Dhdan)	TS2	Strong teratogens	D	
	Methotrexate (Mx)	TS2		X	
	5-Fluorouracil (5FU)	TS1,2		D	
	Busulfan (Bu)	TS1,2	Not classified	D	
	Cytosine arabinoside (CA)	TS1,2		D	
	Hydroxyurea (Hy)	TS1,2		D	
	(<i>trans</i>) Retinoic acid (RA)	TS1,2		X	
	Thalidomide (Thal)	TS1,2		X	
	Valproic acid (VPA)	TS1,2		D	
	Amiodarone (Am) (Blind 3)	TS2		D	
	Rifampicin (Ri) (Blind 4)	TS2	Not classified	C	
	Carbamazepine (Car) (Blind 5)	TS2		C	
	Accutane (Ac) (Blind 6)	TS2		X	
Cyclophosphamide (Cyc) (Blind 7)	TS2	D			

Shading was provided as a visual aid for teratogenicity classification. The darker the shading, the stronger the teratogenicity classification.

with drugs at a concentration equivalent to their published serum circulating therapeutic dosages, typically at the maximum concentration (see Table 2). Dosing was performed on hES cells in 6-well plates in triplicate (“dosed cells”), i.e. three wells per plate. The plates were dosed in triplicate, so there were a total of nine dosed wells. In parallel, there were nine wells in which hES cells were cultured with mTeSR1 containing no drug (“control cells”), and three wells containing Matrigel with mTeSR1 media without hES cells that served as media controls. Lastly, three wells of dosed media controls were prepared, containing Matrigel, mTeSR1 and drug, but no hES cells. These dosed media and media controls provided only background mass spectral data and were not used for fold change calculations other than to determine if a feature was secreted. On the first day of dosage, the determined concentration of drug was dissolved in mTeSR1, and then 2.5 mL of this solution was added to each dosed well of hES cells. Each day, for 4 days, the media was removed and new dosed media containing the drug of interest was added. On the fourth day, the media was removed and added to acetonitrile to make a 40% acetonitrile solution, as outlined in the Sample preparation section below.

hES cell viability assays. Cell viability was assessed in response to exposure to a subset of the test compounds using the MultiTox-Fluor Assay (Promega, Madison, WI), which simultaneously measures cell viability and cytotoxicity. WA09 hES cells were seeded at a density of 250,000 cells/well in a 96-well plate, comparable to the cell density in the 6-well plates used in the metabolomics experiments. Cells were fed with dosed media daily, for 4 days. On the fourth day, spent media was removed, 100 µL of fresh media was added along with 100 µL of

Table 2
Chemical compound dosages and references.

Stemina Classification	Compound	Dosage (µg/mL)	Reference	
Non-teratogens	Ascorbic acid	7.92	(Levine et al., 1996)	
	Isoniazid	2.6	(Seth et al., 1994)	
	Penicillin G	17	(Drug package insert)	
	Folic acid	0.0003	(Kalmbach et al., 2008)	
	Levothyroxine	2.5	(Escobar et al., 2005)	
	Retinol (Blind 1)	0.286	(Gazala et al., 2002)	
	Doxylamine (Blind 2)	0.103	(Friedman et al., 1989)	
	Thiamine (Blind 8)	0.023	(Thornalley et al., 2007)	
	Aspirin	1.805	(Blacklock et al., 2001)	
	Caffeine	4.4	(Benowitz, 1990)	
	Diphenhydramine	0.0353	(Brocks, 1999)	
	Teratogens	Dexamethasone	5.0	(Toth et al., 1999)
		Diphenylhydantoin	13.8	(Schmidt and Kupferberg, 1975)
		Methotrexate	0.0909	(Shoda et al., 2007)
		5-Fluorouracil	27.3	(Muller et al., 1997)
Busulfan		49.6	(Zwavingl et al., 2005)	
Cytosine arabinoside		0.219	(Weinstein et al., 1982)	
Hydroxyurea		10.26	(Gwilt and Tracewell, 1998)	
(trans) Retinoic acid		0.347	(Muindi et al., 1992)	
Thalidomide		0.062	(Torano et al., 1999)	
Valproic acid		166.4	(Bowden, 2007)	
Amiodarone (Blind 3)		0.9	(Drug package insert)	
Rifampicin (Blind 4)		8.0	(Bowman, 1974)	
Carbamazepine (Blind 5)		7.09	(Burstein et al., 2000)	
Accutane (Blind 6)		0.412	(Khoo et al., 1982)	
Cyclophosphamide (Blind 7)		154	(Chen et al., 1997)	

the MultiTox-Fluor reagent. The plate was incubated at 37 °C, 5% CO₂ for 30 min and measured at an excitation wavelength of 490 nm to measure viability and at an excitation wavelength of 485 nm to measure cytotoxicity. The ratios of live to dead cells were normalized to the control cells (no treatment) in order to report relative cell viability.

Sample preparation

For each type of sample – dosed cells, control cells, media and dosed media, 2.5 mL of spent media from each well was added to 1.67 mL acetonitrile (ACN) to make a 40% acetonitrile solution. The acetonitrile acts to “quench” the spent media sample, slowing or halting many metabolic processes and aiding in precipitation of cellular proteins. Samples were either stored at –80 °C for later analysis, or for immediate analysis, 250 µL of the quenched solution was mixed with 250 µL of water, to a final concentration of 20% ACN, then added to a 3 kDa molecular weight cut-off filter spin column (Microcon YM-3 Centrifugal Filter, Millipore, Billerica, MA). Each sample was then centrifuged in an IEC CL31R Multispeed Centrifuge (Thermo Scientific, Waltham, MA) at 13,000×g at 4 °C for 200 min. Following centrifugation, the flow-through was saved then dried for several hours in a Savant High Capacity Speedvac Plus Concentrator. The concentrated sample was then dissolved in 50 µL of 0.1% formic acid prior to LC–MS analysis.

Mass spectrometry

Mass spectrometry was performed using an LC–MS system consisting of a 1200SL HPLC system interfaced with a G6520AA QTOF high-resolution mass spectrometer (Agilent Technologies – Wilmington, DE) capable of exact mass MS and MS–MS data acquisition. In order to facilitate separation of small molecules with a wide range of polarity and to allow increased retention of hydrophilic species, Hydrophilic Interaction Liquid Chromatography (HILIC) (Alpert, 1990) was

employed. A Luna HILIC column (Phenomenex, Torrance, CA) with dimensions 3×100 mm and 3 µm particle size was used and maintained at 30 °C. 5 µL of each sample was injected. The running order of the samples was randomized. Data acquisition time for each sample was 30 min at a flow rate of 0.5 mL/min, with 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B). A sample solvent blank (0.1% formic acid) was run after every 10 samples. Electrospray ionization was employed using a dual ESI source. Spectra were internally mass calibrated in real time using a reference mass solution containing two known reference compounds (HP-921 (Hexakis(1H, 1H, 3H-tetrafluoropropoxy)phosphazine) and 7H-purine), bracketing the mass range of interest at *m/z* 922.00980 and 121.0509 respectively in positive-ion electrospray. Negative-ion spectra were internally calibrated using the TFA (trifluoroacetic acid) anion at *m/z* 112.9856 and the TFA adduct of HP-921 at *m/z* 1033.9881. This calibration mixture was continuously delivered by an Agilent isocratic pump into the electrospray ion source at approximately 0.01 mL/min. For MS experiments, the mass range of the instrument was set to 100–1700 Da. Data acquisition was performed with MassHunter Acquisition software using high-resolution exact mass conditions and each set of samples was run first under ESI positive polarity then under ESI negative polarity conditions. The mass resolution of the instrument (*m/z* dependent on TOF systems) was approximately 3000–12,000. MS–MS experiments were performed as the metabolite validation primary procedure. Targeted MS–MS data acquisition was performed where the *m/z* value of the molecular precursor ion of interest was input into the acquisition software along with a value for CID collision energy and expected retention time range, then the samples were run under identical LC conditions.

Mass spectral data statistical preprocessing. Following LC–MS data acquisition, chromatograms were inspected for reproducibility where the Total Ion Chromatograms (TICs) for each acquisition were graphically overlaid in the Qualitative Analysis software (Agilent Technologies), then datafiles that exhibited TICs that showed significant deviation (>+/-20 s for media peaks) or were missing data or showed extraneous peaks were removed from further processing and analysis. The LC–MS data were then used to create mass features that correspond to molecules detected across the different LC–MS runs. Mass features were extracted from the LC–MS data using the MassHunter Qualitative Analysis software. The following criteria were used as general guidelines, however some flexibility and optimization was needed. The *m/z* values within the range of 75–1500, with a charge of +1 or –1, and a centroid height greater than 1000 were used to generate mass features. The *m/z* peaks that pass these criteria were used to fit isotope and adduct (Na⁺, K⁺, and NH₄⁺) patterns corresponding to individual molecules, and to establish the abundance of each mass feature. The abundance was calculated by MassHunter software as the sum of the isotopic and adduct peaks that correspond to a single molecular feature. After data deconvolution, mass features showing at least two ions (e.g. (M+H)⁺ and (M+H)⁺+1 or (M+H)⁺ and (M+Na)⁺) and an abundance value greater than 50,000 for positive-ion mode data and 10,000 in negative-ion mode data were included in the data set used for binning of the mass features.

Following feature creation by MassHunter, the data was further preprocessed by MassProfiler (Agilent Technologies) software which aligns mass features across multiple LC–MS data files. Mass features were generated for data from all samples for each drug treatment experiment (dosed cells, control cells, media and dosed media samples) using the default alignment settings in MassProfiler with the requirement that a feature be present in at least 80% of the samples in one treatment. The mass feature datasets for each drug treatment experiment were further processed in a global manner using custom analysis scripts executed in the R statistical software environment.

Files for each drug experiment were binned using an algorithm based on both exact mass and retention time in order to consider a mass feature the same across different LC-MS runs. The binning criteria is based on both a sliding mass difference scale that allows for larger mass differences at lower molecular weights and a constant retention time window based on the reproducibility of the chromatography. Masses were ordered and considered to be the same feature if a mass under 175 Da differs by less than 18 ppm from the previous mass, while masses 176–300 Da were binned by 12 ppm and 10 ppm when over 300 Da. These mass bins were ordered by retention time and if a difference in retention of the previous feature was less than 12 s it was considered to be the same feature across LC-MS runs. The binning process is used to create unique compound identities (cpdID) that are assumed to represent a single small molecule. If multiple mass features appeared to fall into the same bin their abundances were averaged.

Definition and determination of “secreted” metabolites. The stem cell growth media (mTeSR) represents a major factor in the experimental system, in that it contributes many peaks to mass spectra. Therefore, in addition to the primary metabolomic analysis focus of detecting fold changes in metabolites by comparison of the spent media from dosed cells and control cells, some metabolites were also defined as being “secreted” by determining mass feature bins which were present at significant levels above the media alone. Mass feature bins present solely in the presence of cells (not detected in media) or with average abundance levels at least 30% different than uncultured media were defined to be “secreted” metabolites.

Validation of small molecule metabolites. Validation experiments were conducted to validate mass features that had been annotated as putative metabolites by acquiring LC-MS-MS data for reference compounds then comparing the spectra to the LC-MS-MS data of the putative metabolites. Reference compounds were purchased and dissolved in mTeSR media and prepared in exactly the same manner (described above) as the samples of spent media from the cells, including the addition of acetonitrile, Centricon centrifugal filtration, drying then dissolution in formic acid prior to LC-MS-MS analysis. The same preparation procedure was followed to ensure chromatographic reproducibility, partially because some reference standards that had been dissolved in a solvent other than mTeSR (such as 0.1% formic acid) showed significant retention time variation. In order to validate the identity of the metabolite, several comparative criteria were required to be met: 1) the retention time of the metabolite detected in the cell media must be within + or – 20 s of the reference standard. 2) The measured molecular exact mass of the metabolite must be within 10 ppm of the known mass of the reference compound. 3) The MS-MS fragmentation spectra of the metabolite was required to be a reasonable match with that of the reference standard where the exact m/z values of the fragment (product) ions must be within 20 ppm of those of the reference standard and the relative abundances of the product ions must be similar. In addition, if published MS-MS spectra were available, these spectra must also be a reasonable match using the same criteria.

Feature set used for random forest modeling. The dataset utilized for random forest modeling was a subset of high quality reproducible mass features. Mass features were selected if they had values present in at least 75% of all drug treatment experiments (blind and known drugs). This list of mass features was then filtered against a list of known contaminant molecules such as HEPEs and PEG related compounds and their numerous adducts to facilitate removal of mass features of non-biological interest. To ensure that only bins with the highest accuracy were used in data analysis, mass features with poor binning or grouping characteristics were removed by filtering the mass feature bins based on mass error distribution and retention

time criteria. Bins were first filtered according to their error distributions where mass feature bins are removed from analysis that show greater than 30 ppm difference between the minimum and maximum mass when compared to the mean mass of the bin. Bins that have retention time differences greater than 1 min and 30 s between the minimum and maximum retention time were also removed from analysis because they represent poor quality feature bins due to (for example) isobaric species with overlapping elution, ions present in the solvent, or other contaminant ions. Bins were first filtered according to their error distributions where mass feature bins are removed from analysis that show greater than 25 ppm difference between the minimum and maximum mass when compared to the mean mass of the bin. Because some mass feature bins may contain only a small number of measurements due to errors associated with integration of background peaks and/or missing isotopes or adducts, these bins that are not highly reproducible were removed. A bin must also contain no more than 33% missing values for given experimental factors or it was removed from analysis. Bins were also removed if they contained more than 50% of the expected number of peaks based on sample count because these bins may represent co-eluting isobaric species. After mass feature bins were selected, the data was normalized to column medians and row median.

The abundance values were then log-base-two transformed and the median value of each treatment (dosed cells and control cells) within each experiment (different drugs) was used. The data was then normalized by control for each drug treatment experiment. The resulting median log fold change values were used as the input data values for the random forest modeling. Missing median log fold change data was replaced with a zero. The remaining positive and negative ESI mode mass features were combined, creating a dataset with 144 mass features used for modeling.

Random forest modeling. Random forest (Brieman, 2001) was used to create a classification model in order to predict teratogenicity and non-teratogenicity using the median fold change for each specific mass feature in dosed stem cells versus the associated intra-experimental control (undosed) stem cells. The random forest models consisted of 1000 decision trees with the square root of the number variables (features) sampled at each split with replacement based on bagging 1/3 of the samples. Final prediction from the RF classifier on the blinded drugs (i.e. drug treatment experiments whose true identifications were coded and unknown to the statistician) was based on the majority vote of the ensemble of trees. Feature selection by variable importance was performed by selecting mass features with a mean decrease in accuracy greater than 0.5. Random forest based analysis was executed using the R randomForest library. Model metrics were calculated based on the resulting random forest confusion matrix or the predictions of blinded drugs using the methods outlined in (Genschow et al., 2000).

Results

Stem cell dosing

A training set of established teratogens and non-teratogens (Table 1) was used to dose hES cells. This training set consisted of a collection of chemical standards that includes compounds that had been previously used in multicenter efforts aimed at developing and validating novel alternatives to predict developmental toxicity, such as the EST, proposed by the ECVAM agency. Since it is our goal to develop a more predictive *in vitro* assay than those currently available, and to further identify biomarkers that are specific to humans, rather than to rodents or other non-human biological systems, we replicated the ECVAM test set in this study. Additional drugs were included in the training set to increase the number of the non-teratogen chemicals, as well as to supplement the strong teratogens.

Compounds were dosed at concentrations corresponding to their circulating dose rather than IC₅₀ or EC₅₀ dose levels (Table 2). Dosing was done at the circulating maternal dose as published in the literature in an effort to recapitulate the exposure level to the developing human embryo *in vivo* and the toxic effect on human development rather than creating a model which measures toxic effect on hES cells in culture. It is noteworthy to mention that the substances employed in this screen (the ECVAM test set) exert their developmental toxicity in a manner that is independent of maternal metabolism.

Teratogen classification for the random forest model

The classification of teratogenicity in previously published animal and cell culture models of developmental toxicity were trained using three different classes, non-teratogens, weak/moderate teratogens, and strong teratogens, based largely on embryotoxicity outcomes and developmental abnormalities observed in animal models (Chapin et al., 2008; Marx-Stoelting et al., 2009). We decided to take a modified approach to compound classification since there are many species-specific differences in developmental toxicity. We based our compound teratogenicity classification strictly on the observed human risk associated with each chemical. These criteria of observed human teratogenicity risk led to a model with two categories of toxicity – teratogen or non-teratogen, which accurately reflects the ultimate intended outcome of the predictive model. This also reduces the technical challenges associated with attempting to determine the potency of teratogens based on distantly related species. We believe a focused classification schema (teratogens versus non-teratogens) leads to a more robust and predictive metabolic model of human developmental toxicity given the limited availability of reliable, quantitative data of human risk associated with exposure to weak or moderate teratogens.

Cell viability assays

In addition to determining teratogenicity by molecular endpoints using metabolomics, cell viability was tested for a subset of the drugs in order to determine if a correlation exists between cell death and compound teratogenicity. In particular, we were concerned that the metabolic endpoints may be strongly correlated with cell death rather than, or in addition to, developmental toxicity. This was particularly true when dosing with the antineoplastic drugs cytosine arabinoside and 5-fluorouracil often resulted in the most profound changes in many metabolites. Cell viability data (Fig. 1) showed no discernable

correlation between teratogenicity and cell death relative to control cells.

Thus, therapeutic concentrations of teratogens are not correlated with cell death in a significant manner, despite the evidence of statistically significant metabolomic changes. This finding suggests that metabolomics has a lower threshold, or increased sensitivity to detect molecular changes associated with developmental toxicity and specific biomarkers in comparison to standard cell death assays, which should provide a more predictive and sensitive screen for developmental toxicity.

Random forest model results

The machine learning method of random forests creates a committee of decision trees by random variable selection from a bootstrapped training set to generate a predictive model based on majority vote of the ensemble of trees. We chose random forest modeling because it is applicable to noisy and high dimensional data that exists in metabolomics datasets, does not require pre-selection of features, is resistant to over fitting, and feature selection produces a small cohort of features with high predictive accuracy (Diaz and Alvarez, 2006). A random forest model was trained using a filtered dataset consisting of reproducibly measured mass features from both ESI polarities. Fold changes (i.e. changes in abundance for a specific mass feature in dosed cells versus its associated intra-experimental control cells), were used as the variables to predict the teratogenicity of drugs.

The initial training set (TS1) contained 144 mass features resulting from exposure of hES cells to 7 teratogens and 5 non-teratogens (see Table 1); these mass features served as the basis for the model that was applied to predict the teratogenicity of chemical compounds in the blind studies. This model was able to correctly predict the teratogenicity of 7 of 8 blinded drug treatments, with a specificity of 100%, sensitivity of 80% and overall accuracy of 88% (Table 3). The random forest model was further refined by integrating the outcomes from the blinded drugs into the model as known classifiers, thereby increasing the number of non-teratogens and teratogens in the model, resulting in a larger training set consisting of 26 drug treatment experiments (TS2, see Table 1). Feature selection based on the variable importance measure mean decrease in accuracy resulted in 18 mass features (Table 4) that were evaluated as a future predictive model. As a result, the overall accuracy of the model was ultimately increased to 92%, i.e. the model was able to correctly predict 24 of the 26 drugs used in the training set. The model was clearly able to differentiate teratogens from non-teratogens into distinct clusters when evaluated by multidimensional scaling (Fig. 2) which reflects

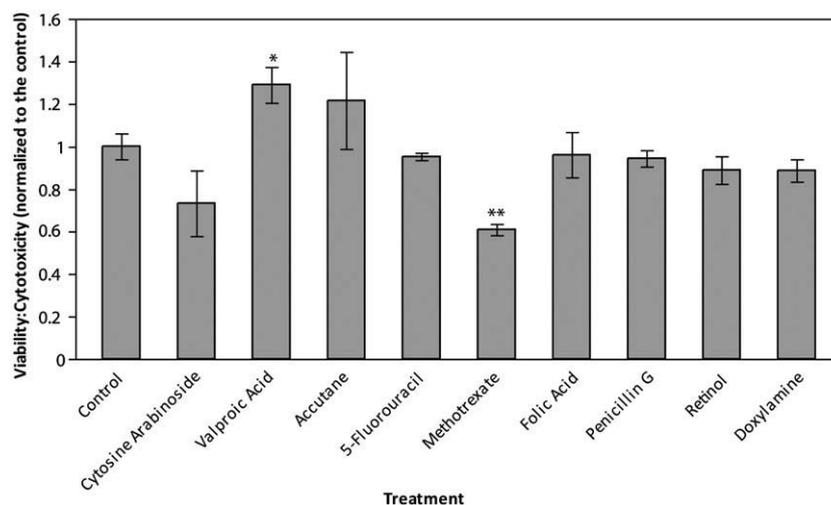


Fig. 1. Cell viability data that has been normalized to control, undosed cells. * $p < 0.05$. ** $p < 0.01$.

Table 3

Results of the blinded study where the teratogenicity was correctly predicted for 7 of 8 drugs using a random forest statistical model.

Blind #	Drug	Actual	Predicted
B1	Retinol	Non	Non
B2	Doxylamine	Non	Non
B3	Amiodarone	Ter	Ter
B4	Rifampicin	Ter	Ter
B5	Carbamazepine	Ter	Ter
B6	Accutane	Ter	Non
B7	Cyclophosphamide	Ter	Ter
B8	Vitamin B1	Non	Non

Bold indicates the incorrect teratogenicity prediction.

clear differences in metabolomics endpoints between treatment classes.

Thus, this model shows great potential for increased accuracy in the prediction of human developmental toxicity in comparison to currently available assays, and that the use of iterative modeling as more experiments are performed is a powerful benefit to the adoption of meaningful metabolic endpoints in a screen. The predictive ability of this model is subject to continuous monitoring in response to additional blinded drug treatments.

Succinic acid was one of the 18 mass features that contributed to the random forest prediction model (PM) and its molecular identity was validated by MS–MS. This metabolite shows significant down regulation in its abundance in response to several teratogens such as carbamazepine, cyclophosphamide, cytosine arabinoside, 5-fluorouracil, methotrexate, and valproic acid.

Results for specific metabolites

Statistically significant differences in the abundance of other specific metabolites were detected in drug-treated and control samples. One such metabolite, asymmetric dimethylarginine (ADMA), exhibited a significant fold decrease in its abundance in response to valproic acid treatment and showed similar changes for the strong teratogens cytosine arabinoside, 5-fluorouracil, hydroxyurea, amiodarone and cyclophosphamide. ADMA was determined to be a secreted compound as defined earlier. ADMA is an inhibitor of nitric oxide synthase (NOS), an enzyme that converts L-arginine to L-citrulline which is necessary for neural tube closure. Valproate is known to cause neural tube defects (DiLiberti et al., 2005) while nitric oxide synthase activity is essential for neural tube closure (Nachmany et al., 2006) and nitric oxide has been shown to induce

Table 4

Features present in the 18 feature random forest model.

Feature ID	Mode	RT (sec)	Mass
pos48	pos	55	103.2278
pos102	pos	80	113.0843
pos563	pos	503	174.2275
pos687	pos	46	189.1261
pos744	pos	70	198.1728
pos1698	pos	42	315.2042
pos1734	pos	597	320.0177
pos1791	pos	506	328.0611
pos2109	pos	498	374.0676
pos2814	pos	655	478.1241
pos3308	pos	82	564.2780
pos3832	pos	92	662.8337
pos4617	pos	83	844.4077
neg71	neg	81	118.0265
neg133	neg	75	132.0779
neg360	neg	438	187.0445
neg414	neg	496	200.0279
neg1121	neg	55	331.0842

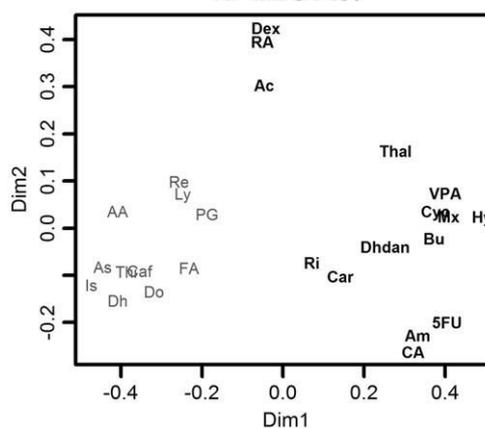
RF MDS Plot

Fig. 2. Multidimensional scaling plot set in two dimensions of the distances based on the proximity matrix from the 18 feature random forest model showing a clear separation of drugs (abbreviations in Table 1) based on teratogenicity. Rifampicin and accutane were misclassified as non-teratogens by the random forest model. Gray = non-teratogen, black = teratogen.

neural tube defects (NTD) in rat embryos. The novel alterations in the secretion of ADMA, detected here, suggest that it may be an appropriate candidate biomarker for neural tube defects. Arginine levels were also monitored in our data and usually showed opposite fold changes to those of ADMA in response to several strong teratogens. To quantify the perturbation of arginine and ADMA in the hES cells as a result of dosing, EICs (Extracted Ion Chromatograms) for these compounds were constructed and their areas were integrated, then the ratio of the resulting areas for controls vs. dosed were compared. These results indicate that the amount of perturbation may be directly related to the teratogenicity of the dosing compound. There are no false negatives resulting from these metrics, and only ascorbic acid and caffeine are false positives for teratogenicity (Table 5).

Other small molecules, including those present in the mTeSR media – gamma aminobutyric acid (GABA), isoleucine, aspartic acid, glutamic acid, histidine and arginine, and those secreted molecules not present in the mTeSR media – succinic acid and malic acid, showed measureable fold changes. Detected levels of these small molecules in many cases were altered according to the teratogenicity of the test compound and can be correlated to each other on the basis of the biochemical pathways where they serve as intermediates. This is illustrated in Fig. 3.

Discussion

Potential advantages of a hES cell developmental toxicity prediction model

The hES cell assay reported herein has several potential advantages over other standard approaches, namely: 1) alterations to the metabolites in response to a toxicant is a sensitive and quantitative measurement, which enables more objective data-driven decisions. 2) Multiple biochemical pathways can be assessed simultaneously, which reinforces the robustness of the model when applied to drugs with a variety of mechanisms of toxicity. 3) Metabolic endpoints are a measure of functional biochemical pathways that may be rapidly integrated with protein, DNA, and RNA targets for further pathway-based investigation. 4) Since the prediction is based on multiple independent variables, it is possible to detect teratogens exhibiting complex changes in metabolic patterns. 5) The assay is independent of cell death outcomes and was trained on circulating doses known to cause human developmental toxicity, which increases the probability

Table 5

Selected fold change ratios for arginine and asymmetric dimethylarginine (ADMA). EICs for these compounds were integrated, then the fold change of the resulting areas for controls vs. dosed were compared. Smaller fold change ratios (between 0.9 and 1.1) show a good correlation with non-teratogens, while greater changes (<0.9 and >1.1) correlate with teratogens. There are no false negatives for teratogenicity resulting from these metrics and only ascorbic acid and caffeine are false positives.

Stemina classification	Compound	Arg fold change/ADMA fold change	Arg/ADMA prediction	
Non-teratogens	Ascorbic acid	1.28	Ter	
	Aspirin	1.07	Non	
	Caffeine	1.33	Ter	
	Doxylamine (Blind 2)	0.97	Non	
	Isoniazid	0.94	Non	
	Levothyroxine	1.03	Non	
	Penicillin G	0.96	Non	
	Folic acid	1.08	Non	
	Retinol (Blind 1)	1.03	Non	
	Thiamine (Blind 8)	1.00	Non	
	Teratogens	5-Fluorouracil	43.93	Ter
		Methotrexate	2.54	Ter
		Accutane (Blind 6)	0.55	Ter
Amiodarone (Blind 3)		1.64	Ter	
Busulfan		1.12	Ter	
Carbamazepine (Blind 5)		1.12	Ter	
Cyclophosphamide (Blind 7)		1.56	Ter	
Cytosine arabinoside		67.01	Ter	
Hydroxyurea		2.52	Ter	
Retinoic acid		0.48	Ter	
Rifampicin (Blind 4)		0.81	Ter	
Thalidomide		0.85	Ter	
Valproic acid		2.11	Ter	

of finding developmental toxicants that are not just toxic to dividing cells. 6) With additional development, testing and analysis can be made higher throughput, less labor intensive and more automated.

Faithful recapitulation of whole organs has not been demonstrated by differentiation of hES cells. Thus, despite producing functional human cell types, their ability to wholly mimic human development is incomplete. For example, the disruption of pathways that are specific to gastrulation through organogenesis cannot be measured *in vitro*. We acknowledge that the hES cell model will provide, at best, an incomplete portrait of toxicant response. However, no single model has such potential. Factors such as maternal or hepatic metabolism and bioavailability will not come into play. For ethical reasons, toxicants cannot be tested in humans and especially in human embryos. Thus recognizing the principal limitation of hES cells, i.e. the absence of organogenesis *in vitro*, the pluripotency of hES cells still

offers a useful surrogate assay to evaluate toxicant risk; as a result of pluripotency, hES cells undergo key events of morphogenesis such as embryoid body formation, cardiogenesis, and neurogenesis. In fact, the duration of neuronal and glial differentiation from hES-derived neural precursors is similar to lineage allocation *in vivo* (Zhang, 2006). It is plausible to speculate that this system will enable the discovery of tissue-specific biomarkers of developmental toxicity, following exposure of hES-differentiated cells (neural, cardiac and osteogenic precursors, for example) to known chemical disruptors of human development.

Comparison of hES cell developmental toxicity prediction model to other models

Despite the evidence for the application of hES cells and metabolomics to predict developmental toxicity reported herein, it is noteworthy to mention that *in vitro* systems cannot fully recapitulate or account for key events that may significantly contribute to disruption of normal human development by exogenous chemicals. Absorption, distribution, metabolism and excretion (ADME) cannot be properly modeled in an *in vitro* system, which may make it difficult to predict how a substance of unknown toxicity will behave *in vivo*. In addition, the hES cell model cannot fully emulate teratogenicity *in utero*, which ultimately results also from systemic effects, in particular fetal-maternal interactions and the complex process of organogenesis. Nonetheless, given the physiological relevance of hES cells to human development, developmental toxicity testing in cells derived from human embryos may generate more reliable *in vitro* prediction endpoints than those currently available through the use of animal models or other *in vitro* non-human assays such as zebra fish models, the EST, and whole embryo culture (WEC).

The hES cell model has important biological features in comparison to zebrafish assay systems. First, it is a human system, providing species specificity to predict human outcomes. Zebrafish developmental and biochemical pathways can be quite distinct from those that are critical to human development, for example the absence of placentation and pulmonary differentiation and development, as well as different mechanisms for cardiogenesis. Moreover, the screening throughput of zebrafish assays is somewhat limited due to the high degree of developmental defects associated with small well size (Selderslaghs et al., 2009). The fish are also sensitive to very low concentrations of DMSO, where levels greater than 0.25% cause increased deformities. The observation of a specific defect by visual inspection of changes in morphology can also be highly subjective, while perturbation to the abundance of small molecule metabolites in

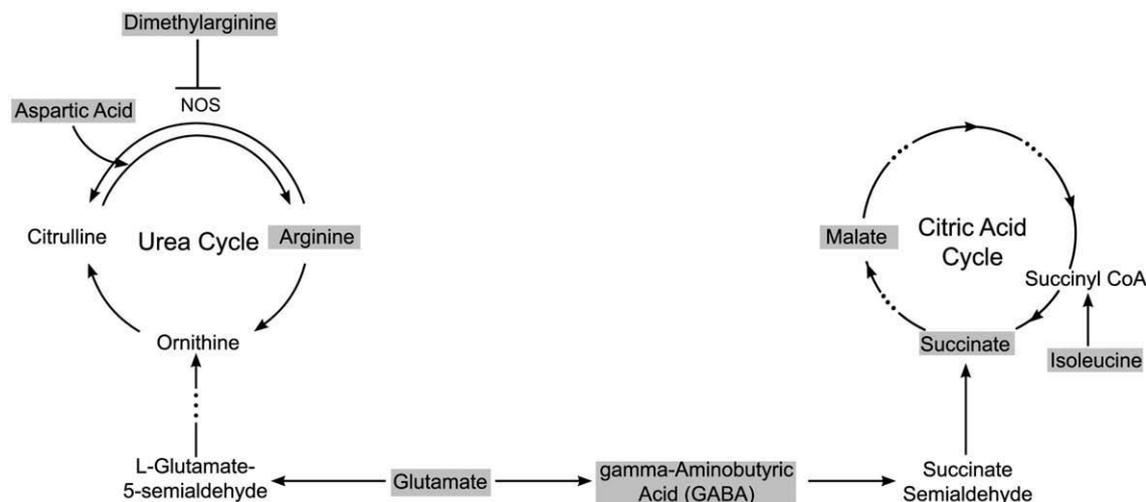


Fig. 3. Illustration of the metabolic pathway relationships between the metabolites found in this study.

the hES assay is a quantitative endpoint measured by a highly sensitive and specific analytical technique (LC–ESI–QTOF–MS).

The European Centre for the Validation of Alternative Methods (ECVAM) has approved the mouse embryonic stem cell test (EST) as a teratogenicity screen. This test accurately predicted the teratogenicity of 78% of 20 compounds in a test panel (Genschow et al., 2002). However, the biological mechanisms of the EST are still poorly understood, the test is not specific to human response and the endpoint (number of beating heart cells differentiated from embryoid bodies) is somewhat subjective (Adler et al., 2008).

The EST predictive model is strongly correlated with cytotoxicity, given that two EST variables result from the IC₅₀ concentrations observed in fibroblasts compared to mES cells. These variables make the assumption that developmental toxicants cause cell death at lower concentrations in embryonic cells compared to the “adult” fibroblast cells, which may not be valid for many mechanisms of toxicity (for example – Thalidomide). The dose required to reach an IC₅₀ may also be much higher than the typical circulating dose or that which may be encountered by the fetus *in utero* leading to large numbers of false positives. It is also likely that changes in cell viability may be observed *in vitro* which will not occur *in vivo*. Therefore, in comparison to teratogenicity determinations reported for the EST, which measures cytotoxicity and the ability of chemicals to disrupt proper differentiation of mES cells into cardiomyocytes, the overall potential accuracy of the hES cell assay reported here may exceed the ability of the EST to deliver objective measures of human developmental toxicity.

The initial results indicate that the overall accuracy of the current hES cell metabolomics based developmental toxicity assay is 88%, compared to 83% for the modified version of the EST (Paquette et al., 2008), and 78% for the original EST method and the hES cell based assay correctly classifies thalidomide as a teratogen while the EST does not (Nieden et al., 2001). So far, the hES cell model appears to be considerably more predictive than either WEC or micro mass (MM) (Table 6). Further, the molecular endpoints measured in the hES cell and metabolomics-based model offer an opportunity to understand the mechanisms of developmental toxicity in an all-human system.

Mechanistic pathways of developmental toxicity

Altogether, when comparing dosed to undosed cells, metabolomics of hES cells detected statistically significant changes to levels of multiple small molecule metabolites which play a key role in cellular physiology and human development. Several of the small molecules that showed significant and reproducible fold changes between dosed and undosed cells, are contained in the mTeSR cell media. These fold changes indicate that cell metabolism of these compounds has been directly altered by the drug dose and that the metabolic equilibrium or flux balance has been disturbed, therefore these small molecules may also serve as biomarkers for teratogenicity. Several of these candidate biomarkers were further validated by MS–MS in order to confirm their chemical identity. Significantly, despite the unsupervised nature of the analysis, many of these significant and validated

small molecule metabolites participate in pathways that had been previously suggested to underlie developmental toxicity albeit not in cells derived from human embryos. As reported in the Results section, ADMA, an inhibitor of Nitric oxide (NO) metabolism, exhibited significant increases in fold changes in response to exposure of hES cells to strong teratogens. NO has been identified as a candidate mechanism for neural tube disorders, and NO is essential for normal axial development (Alexander et al., 2007). It has been previously shown that both an excess as well as a deficiency of NO can be embryotoxic (Lee and Juchau, 2005). The present study is the first time that two human intermediates in this pathway, arginine and ADMA (Table 5) were measured and exhibited statistically significant changes in response to several known disruptors of human development.

As shown in Fig. 3, metabolites showing statistically significant fold changes include aspartic acid, ADMA, and arginine, which are all components of the urea cycle. This cycle facilitates the removal of dangerous ammonia through conversion of it to urea, which is excreted from the body. Succinic acid, isoleucine, and malate are part of the citric acid cycle, which produces energy for cellular function. Both pathways are linked by glutamate and GABA, which in turn has a critical role in neuronal physiology.

Certain reactions in the urea cycle take place in the mitochondria, while the Krebs cycle is active in the mitochondria in its entirety. Perturbations to the urea cycle can result in excess ammonia, which, among a vast array of pathological effects, has been correlated to newborn deaths (Summar, 2001). Interruption of citric acid cycle reactions compromises cellular energy metabolism with direct detrimental effects to cellular viability.

In our studies, increased concentrations of GABA were detected in the secretome of hES cells dosed with busulfan, among other teratogens. Dysfunctions in GABA, underlie well established neurological disorders such as epilepsy, language delay, and neurodevelopmental impairment, among others (Pearl and Bigson, 2004). The neurodevelopmental toxicity of busulfan has been previously reported in humans; specifically *in utero* exposure to busulfan led to a spinal birth defect due to insufficient neural fold development, although the mechanism was not defined (Abramovici et al., 2005). The metabolomics results presented here suggest that busulfan affects GABA levels in the developing embryo, which in turn may underlie neural developmental disruption. These examples illustrate how metabolomics can unravel mechanistic pathways of developmental toxicity through direct analysis of secreted metabolites from hES cells dosed with known teratogens. In doing so, it is quite possible to model the potential for developmental toxicity of new drugs screened in preclinical development with a high degree of predictability while providing information about the mechanisms of toxicity. Further studies may allow classification of compounds into subgroups of developmental toxicity such as neural developmental disruptors or those likely to cause structural malformations.

Other key small molecules that showed fold changes, as reported in the results section, share the same chemical pathway, namely GABA and glutamic acid. GABA is the principal inhibitory neurotransmitter in the brain. Glutamate dysregulation has the potential to severely compromise neurogenesis, possibly contributing to cell death in specific regions of the brain (reviewed in Bauman, 1998). Specifically, glutamate is vital for programmed cell death from development until 3 years of age. Not only does the metabolite glutamate regulate neuronal survival or death, but it also plays a critical role in cognition, learning and memory (Tashiro et al., 2006). Glutamate and GABA are also known modulators of neuronal migration during development (Lujan et al., 2005); hence concomitant dysregulation of glutamate and GABA metabolism may provide an important mechanism for human developmental toxicity. Surprisingly, other small molecules reported herein, such as succinic acid, are likely to play synergistic roles with glutamic acid and GABA in the mechanism of teratogen-

Table 6
Accuracy of developmental toxicity models.

Model	# Drugs	Accuracy
Zebra fish (McGrath 2008)	12	91
devTOX (hES cells)	8	88
EST (Paquette 2008)	63	83
WEC (Genschow 2002)	14	80
EST (Genschow 2002)	20	78
Zebra fish embryos (Chapin 2008)	18	72
MM (Genschow 2002)	20	70
WEC (Genschow 2002)	14	68

Bold indicates the toxicity model described in this paper.

induced toxicity, given that simultaneous changes to rate-limiting enzymes in both pathways (GABA-transaminase and succinic semialdehyde dehydrogenase) are present in certain neuropsychiatric disorders, such as succinic semialdehyde dehydrogenase deficiency or GABA aciduria (reviewed in Pearl et al., 2007). Although this syndrome is inherited, in contrast to the environmental nature of developmental toxicity, it becomes even more striking that valproate has been shown to aggravate symptoms in these patients, through further detriment to GABA and succinic acid metabolism (Shinka et al., 2003), which is a direct indication of the potential of this hES cell based developmental toxicity screen to elucidate biologically meaningful mechanisms of compound toxicity.

Future directions

In addition to the metabolites described herein, a virtual library containing all the biomarkers discovered in this study was established. This library provides a repository of human biomarkers useful in assessing developmental toxicity, not only of pharmaceutical agents, but also of other chemicals, the latter subject to increased attention from regulatory directives, particularly REACH, (Registration, Evaluation, Authorisation and Restriction of Chemical substances). REACH is a new European Community Regulation on chemicals and their safe use (EC 1907/2006). Future studies will integrate a larger number of pharmaceutical compounds in addition to other chemicals that are known to disrupt human development (such as chlorpyrifos, organophosphates and methylmercury) to further expand the biomarker library and the robustness of metabolomics biomarkers across very diverse collections of chemicals. We are also developing a high-throughput developmental toxicity assay based on metabolomics of hES cells in a 96-well format. This will enable high-throughput screening of chemical collections such as those available at the Molecular Libraries Program (NIH) or NTP (National Toxicology Program, NIEHS). In addition, we are directing our approach toward a more targeted metabolomics technique employing the use of triple quadrupole MS using MRM (Multi Reaction Monitoring) mass spectral analysis for increased speed, sensitivity and specificity for quantitation of metabolites.

While future efforts to improve upon this model will be made, the current study illustrates the ability to utilize hES cells and metabolomics to provide a highly predictive, quantitative, all-human *in vitro* screening method for predicting developmental toxicity of compounds with the additional benefit of providing additional data supporting mechanisms of compound toxicity. Thus, this method has the potential to aid in the prevention of birth defects induced by chemical compounds and to significantly reduce animal testing.

Acknowledgments

The authors are grateful to Ron Zhu and BioInfoRx for the Stemina LIMS and to Agilent Technologies for equipment and support. We also wish to thank Sudeepa Bhattacharyya for her contributions.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.taap.2010.05.007.

References

Abramovici, A., Shaklai, M., Pinkhas, J., 2005. Myeloschisis in a six weeks embryo of a leukemic woman treated with busulfan. *Teratology* 18, 241–245.
Adler, S., Pellizzer, C., Hareng, L., Hartung, T., Bremer, S., 2008. First steps in establishing a developmental toxicity test method based on human embryonic stem cells. *Toxicol. In Vitro* 22, 200–211.

Alexander, P., Chau, L., Tuan, R., 2007. Role of nitric oxide in chick embryonic organogenesis and dysmorphogenesis. *Birth Defects Res. A Clin. Mol. Teratol.* 79, 581–594.
Alpert, A., 1990. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids, and other polar compounds. *J. Chromatogr.* 499, 177–196.
Bauman, K.A., 1998. Neuropathology of infantile autism. *J. Neuropathol. Exp. Neurol.* 57, 645–652.
Benowitz, N.L., 1990. Clinical pharmacology of caffeine. *Annu. Rev. Med.* 41, 277–288.
Blacklock, C.J., Lawrence, J.R., Wiles, D., Malcolm, E.A., Gibson, I.H., Kelly, C.J., Peterson, J. R., 2001. Salicylic acid in the serum of subjects not taking aspirin. Comparison of salicylic acid concentrations in the serum of vegetarians, non-vegetarians, and patients taking low dose aspirin. *J. Clin. Pathol.* 54, 533–555.
Bogdanov, M., Matson, W.R., Wang, L., Matson, T., Saunders-Pullman, R., Bressman, S.S., Beal, M.F., 2008. Metabolomic profiling to develop blood biomarkers for Parkinson's disease. *Brain* 131, 389–396.
Bowden, C.L., 2007. Spectrum of effectiveness of valproate in neuropsychiatry. *Expert Rev. Neurother.* 7, 9–16.
Bowman, G., 1974. Serum concentration and half-life of rifampicin after simultaneous oral administration of aminosalicilic acid or isoniazid. *Eur. J. Clin. Pharmacol.* 7, 217–225.
Brent, R., Holmes, L., 1988. Clinical and basic science from the thalidomide tragedy: what have we learned about the causes of limb defects? *Teratology* 38, 241–251.
Brieman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
Brocks, D.R., 1999. Anticholinergic drugs used in Parkinson's disease: an overlooked class of drugs from a pharmacokinetic perspective. *J. Pharm. Pharm. Sci.* 2, 39–46.
Burstein, A.H., Horton, R.L., Dunn, T., Alfaro, R.M., Piscitelli, S.C., Theodore, W., 2000. Lack of effect of St. John's Wort on carbamazepine pharmacokinetics in healthy volunteers. *Clin. Pharmacol. Ther.* 68, 605–612.
Cezar, G., Quam, J., Smith, A., Rosa, G., Piekarczyk, M., Brown, J., Gage, F., Muotri, A., 2007. Identification of small molecules from human embryonic stem cells using metabolomics. *Stem Cells Dev.* 16, 869–882.
Chapin, R., Augustin-Rauch, K., Beyer, B., Daston, G., Finnell, R., Flynn, T., Hunter, S., Mirkes, P., O'Shea, K., Piersman, A., Sandler, D., Vanparys, P., Maele-Fabry, G.V., 2008. State of the art in developmental toxicity screening methods and a way forward: a meeting report addressing embryonic stem cells, whole embryo culture, and zebrafish. *Birth Defects Res. B Dev. Reprod. Toxicol.* 83, 446–456.
Chen, T., Kennedy, M.J., Anderson, L.W., Kiraly, S.B., Black, K.C., Colvin, O.M., Grochow, L.B., 1997. Nonlinear pharmacokinetics of cyclophosphamide and 4-hydroxycyclophosphamide/aldophosphamide in patients with metastatic breast cancer receiving high-dose chemotherapy followed by autologous bone marrow transplantation. *Drug Metabol. Dispos.* 25, 544–551.
Diaz, U.R., Alvarez, A.S., 2006. Gene selection and classification of microarray data using random forest. *BMC Informatics* 7.
DiLiberti, J., Fardoun, P., Dennis, N., Curry, C., 2005. The fetal valproate syndrome. *Am. J. Med. Genet.* 19, 473–481.
Escobar, H.F., Botella-Carretero, J.I., Escobar del Rey, F., Morreale De Escobar, G., 2005. Treatment of hypothyroidism with combinations of levothyroxine plus liothyroxine. *J. Clin. Endocrinol. Metabol.* 90, 4946–4954.
Finnell, R., 1999. Teratology: general considerations and principles. *J. Allergy Clin. Immunol.* 103, S337–S342.
Friedman, H., Greenblatt, D.J., Scavone, J.M., Burstein, E.S., Ochs, H.R., Harmatz, J.S., Shader, R.I., 1989. Clearance of the antihistamine doxylamine reduced in elderly men but not in elderly women. *Clin. Pharmacokinet.* 16, 312–316.
Gazala, E., Sarov, B., Hershkovitz, E., Edvardson, S., Sklan, D., Katz, M., Friger, M., Gorodischer, R., 2002. Retinol concentration in maternal and cord serum: its relation to birth weight in healthy mother–infant pairs. *Early Hum. Dev.* 70, 19–28.
Genschow, E., Scholz, G., Brown, N., Piersma, A., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Breme, S., Becker, K., Spielmann, H., 2000. Development of prediction models for three *in vitro* embryotoxicity tests in an ECVAM validation study. *In Vitro Mol. Toxicol.* 13, 51–66.
Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N., Piersma, A., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S., Becker, K., 2002. The ECVAM international validation study on *in vitro* embryotoxicity tests: results of the definitive phase and evaluation of prediction models. *European Centre for the Validation of Alternative Methods. Alternatives to Laboratory Animals* 30, 151–176.
Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., Kell, D.B., 2004. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol.* 22, 245–252.
Gwilt, P.R., Tracewell, W.G., 1998. Pharmacokinetics and pharmacodynamics of hydroxyurea. *Clin. Pharmacokinet.* 34, 347–358.
Kalmbach, R.D., Choumenkovitch, S.F., Troen, A.M., D'Agostino, R., Jacques, P.F., Selhub, J., 2008. Circulating folic acid in plasma: relation to folic acid fortification. *Am. J. Clin. Nutr.* 88, 763–768.
Keirstead, H., 2005. Human embryonic stem cell-derived oligodendrocyte progenitor cell transplants remyelinate and restore locomotion after spinal injury. *J. Neurosci.* 25, 4694–4705.
Khoo, K.C., Reik, D., Colburn, W.A., 1982. Pharmacokinetics of isotretinoin following a single oral dose. *J. Clin. Pharmacol.* 22.
Kind, T., Tolstikov, V., Fiehn, O., weiss, R., 2007. A comprehensive urinary metabolomic approach for identifying kidney cancer. *Anal. Biochem.* 363, 185–195.
Lee, Q., Juchau, M., 2005. Dysmorphic effects of nitric oxide (NO) and NO-synthase inhibition: studies with intra-amniotic injections of sodium nitroprusside and N-monomethyl-L-arginine. *Teratology* 49, 452–464.
Levine, M., Conroy-Cantilena, C., Welch, R.W., Washko, P.W., 1996. Vitamin C pharmacokinetics in healthy volunteers: evidence for a recommended allowance. *Proc. Natl. Acad. Sci. U. S. A.* 93, 3704–3709.

- Lujan, R., Shigemoto, R., Lopez-Bendito, G., 2005. Glutamate and GABA receptor signalling in the developing brain. *Neuroscience* 130, 567–580.
- Marx-Stoelting, P., Adriaens, E., Ahr, H., Bremer, S., Garthoff, B., Gelbke, H., Piersma, A., Pellizzer, C., Reuter, U., Rogiers, V., Schenk, B., Schwengberg, S., Seiler, A., Spielmann, H., Steemans, M., Stedman, D., Vanparys, P., Vericat, J., Verwei, M., Water, F.v.d., Weimer, M., Schwarz, M., 2009. A review of the implementation of the embryonic stem cell test (EST). The report and recommendations of an ECVAM/ReProTect Workshop. *Altern. Lab. Anim.* 37, 313–328.
- Miller, M., Stromland, K., 1999. Teratogen update: Thalidomide: a review, with a focus on ocular findings and new potential uses. *Teratology* 60, 306–321.
- Muindi, J.R.F., Frankel, S.R., Huselton, C., DeGrazia, F., Garland, W.A., Young, C.W., Warrell, R.P., 1992. Clinical pharmacology of oral all-trans ra in patients with acute promyelocytic leukemia. *Cancer Res.* 52, 2138–2142.
- Muller, M., Mader, R.M., B.S., 1997. 5-Fluorouracil kinetics in the interstitial tumor space: clinical response in breast cancer patients. *Cancer Res.* 57, 2598–2601.
- Nachmany, A., Gold, V., Tsur, A., Arad, D., Weil, M., 2006. Neural tube closure depends on nitric oxide synthase activity. *J. Neurochem.* 96, 247–253.
- Nieden, N.Z., Ruf, L., Kempka, G., Hildebrand, H., Ahr, H., Nieden, N.Z., Ruf, L., Kempka, G., Hildebrand, H., Ahr, H., 2001. Molecular markers in embryonic stem cells. *Toxicol. in Vitro* 15, 455–461.
- Paquette, J., Kumpf, S., Streck, R., Thomson, J., Chapin, R., Stedman, D., 2008. Assessment of the embryonic stem cell test and application and use in the pharmaceutical industry. *Birth Defects Res. B Dev. Reprod. Toxicol.* 83, 104–111.
- Pearl, P., Bigson, K., 2004. Clinical aspects of the disorder of GABA metabolism in children. *Curr. Opin. Neurol.* 17, 107–113.
- Pearl, P., Taylor, J., Trzcinski, S., Sokohl, A., 2007. The pediatric neurotransmitter disorders. *J. Child Neurol.* 22, 606–616.
- Sabatine, M., Liu, E., Morrow, D., Heller, E., McCarroll, R., Wiegana, R., Berriz, G., Roth, F., Gerszten, R., 2005. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 112, 3868–3875.
- Schmidt, D., Kupferberg, H., 1975. Diphenylhydantoin, phenobarbital, and primidone in saliva, plasma, and cerebrospinal fluid. *Epilepsia* 16, 735–741.
- Selderslaghs, I., Rompay, A., De Coen, W., Witters, H., 2009. Development of a screening assay to identify teratogenic and embryotoxic chemicals using the zebrafish embryo. *Reprod. Toxicol.* 28, 308–320.
- Serkova, N., Zhang, Y., Coatney, J., Hunter, L., Wachs, M., Niemann, C., Mandell, M., 2007. Early detection of graft failure using the blood metabolic profile of a liver recipient. *Transplantation* 83, 517–521.
- Seth, V., Seth, S.D., Beotra, A., Semwal, O.P., D'Monty, V., Mukhopadhy, S., 1994. Isoniazid and acetylisoniazid kinetics in serum and urine in pulmonary primary complex with intermittent regimen. *Indian Pediatr.* 31, 279–285.
- Shanks, N., Greek, R., Greek, J., 2009. Are animal models predictive for humans? *Philos. Ethics Humanit. Med.* 4, 2.
- Shinka, T., Ohfu, M., Hirose, S., Kuhara, T., 2003. Effect of valproic acid on the urinary metabolic profile of a patient with succinic semialdehyde dehydrogenase deficiency. *J. Chromatogr. B* 792, 99–106.
- Shoda, H., Inokuma, S., Yajima, N., Tanaka, Y., Oobayashi, T., Setoguchi, K., 2007. Higher maximal serum concentration of methotrexate predicts the incidence of adverse reactions in Japanese rheumatoid arthritis patients. *Mod. Rheumatol.* 17, 311–316.
- Sreekumar, A., Poisson, L., Rajendiran, T., Khan, A., Cao, Q., Yu, J., Laxman, B., Mehra, R., Lonigro, R., Li, Y., Nyati, M., Ahsan, A., Kalyana-Sundaram, S., Han, B., Cao, X., Byun, J., Omenn, G., Ghosh, D., Pennathur, S., Alexander, D., Verger, A., Shuster, J., Weir, J., Varambally, S., Beecher, C., Chinnaiyan, A., 2009. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 457, 910–914.
- Summar, M., 2001. Current strategies for the management of neonatal urea cycle disorders. *J. Pediatr.* 131, S30–S39.
- Tan, D., Manchester, L., Reiter, R., Plummer, B., Hardies, L., Weintraub, S., Vijayalaxmi, Shepherd, A., 1998. A novel melatonin metabolite, cyclic 3-hydroxymelatonin: a biomarker of in vitro hydroxyl radical generation. *Biochem. Biophys. Res. Commun.* 253, 614–620.
- Tashiro, A., Sandler, V., Toni, N., Zhao, C., Gage, F., 2006. NMDA-receptor-mediated, cell-specific integration of new neurons in adult dentate gyrus. *Nature* 442, 929–933.
- Thomson, J., Itskovitz-Eldor, J., Shapiro, S., Waknitz, M., Swiergiel, J., Marshall, V., Jones, J., 1998. Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145–1147.
- Thornalley, P.J., Babaei-Jadidi, R., Ali, H.A., Rabbani, N., Antonysunil, A., Larken, J., Ahmed, A., Rayman, G., Bodmer, C.W., 2007. High prevalence of low plasma thiamine concentration in diabetes linked to a marker of vascular disease. *Diabetologia* 50, 2164–2170.
- Toms, D., 1962. Thalidomide and congenital abnormalities. *Lancet* 2, 400.
- Torano, J.S., Verbon, A., Guchelvar, H.J., 1999. Quantitative determination of thalidomide in human serum with high-performance liquid chromatography using protein precipitation with trichloroacetic acid and ultraviolet detection. *J. Chromatogr. B* 734, 203–210.
- Toth, G.G., Kloosterman, C., Uges, D.R.A., Jonkman, M.F., 1999. Pharmacokinetics of high-dose oral and intravenous dexamethasone (1999). *Ther. Drug Monit.* 21, 532–535.
- Vriend, J., Alexiuk, N., 1996. Effects of valproate on amino acid and monoamine concentration in striatum of audiogenic seizure-prone balb/c mice. *Mol. Chem. Neurobiol.* 27, 307–324.
- Weinstein, H.J., Griffin, T.W., Feeney, J., Cohen, H.J., Propper, R.D., Sallan, D.E., 1982. Pharmacokinetics of continuous intravenous and subcutaneous infusions of cytosine arabinoside. *Blood* 59, 1351–1353.
- Zhang, S.C., 2006. Neural subtype specification from embryonic stem cells. *Brain Pathol.* 16, 132–142.
- Zhang, S.W.M., Duncan, I.D., Brustle, O., Thomson, J.A., 2001. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nat. Biotechnol.* 19, 1129–1133.
- Zwaving, I., Bredium, R.G.M., Cremers, S.C.L.M., 2005. Intravenous Busulfan in children prior to stem cell transplantation: study of pharmacokinetics. *Bone Marrow Transplant.* 35, 17–23.