Innovative approaches in the embryonic stem cell test (EST)

Peter T. Theunissen¹², Aldert H. Piersma¹³

¹Laboratory for Health Protection Research, National Institute for Public Health and the Environment, RIVM, Bilthoven, The Netherlands, ²Department of Toxicogenomics, Maastricht University, Maastricht, The Netherlands, ³Institute for Risk Assessment Sciences, Veterinary Faculty, Utrecht University, Utrecht, The Netherlands

TABLE OF CONTENTS

1. Abstract
2. Development of the embryonic stem cell test (EST)
   2.1. Definition of the EST
   2.2. ECVAM validation of the EST
   2.3. EST: Post validation
   2.4. Defining the applicability domain
   2.5. The applicability domain of compound classes for the EST
   2.6. Implementation of the EST within testing strategies
3. Quantitative endpoints for the EST
   3.1. Alternative EST endpoints
   3.2. Optimization of the classical EST
   3.3. Molecular endpoints in the EST
   3.4. Omics technologies as endpoints for the EST
4. Alternative differentiation models
   4.1. Neural differentiation EST
   4.2. Osteoblast differentiation EST
   4.3. Metabolism in the EST
   4.4. Human EST
5. Future Developments
6. References

1. ABSTRACT

The embryonic stem cell test (EST) is a high-throughput in vitro screening assay for developmental toxicity free of animal use. The EST uses the ability of murine embryonic stem cells to differentiate into the mesodermal cardiac lineage in combination with two cytotoxicity test systems. Validation of the EST showed that the test system is very promising as an alternative method to animal testing, however to optimize predictability and increase knowledge on the applicability domain of the EST, improvements to the method were proposed and studied. In this review we discuss the first definition of the EST followed by the innovative approaches which have been proposed to increase the predictivity of the EST, including implementation of molecular endpoints in the EST, such as omics technologies and the addition of alternative differentiation models to the testing paradigm, such as neural and osteoblast differentiation and the use of human stem cells. These efforts to improve the EST increase the value of embryonic stem cells used as in vitro systems to predict developmental toxicity.

2. DEVELOPMENT OF THE EMBRYONIC STEM CELL TEST (EST)

2.1. Definition of the EST

Throughout life, humans are exposed to numerous different compounds and pharmaceuticals, which may cause a health risk. To detect the potential risk of compounds and pharmaceuticals, a range of accepted in vivo test models, assigned by the OECD (Organization of Economic Co-operation and Development) and ICH (The International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use) is used. Currently, compound and pharmaceutical toxicity testing depends primarily on the use of test animals. As a result of high costs of time consuming animal test protocols and the increasing public ethical awareness to reduce animal testing, it is highly desirable to develop high-throughput alternative screening methods for toxicity testing. Among the different areas of toxicological testing, developmental toxicology requires a high number of test animals for compound toxicity testing within REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) and is required to be
Innovative approaches in the embryonic stem cell test

Figure 1. Schematic overview of the EST culture method

evaluated for pharmaceuticals (1, 2). To reduce the number of animals needed for developmental toxicity testing and screening, a range of *in vitro* test systems have been developed over the past decades, including the whole embryo culture (WEC), the micromass test and the zebrafish embryo test (3-6). However, these systems all require the use of test animals, and are therefore not completely animal free alternatives. A potential high-throughput *in vitro* screening assay for developmental toxicity free of animal use is the embryonic stem cell test (EST), first described in 1997 by Spielmann et al. (7). This test uses the ability of pluripotent murine embryonic stem cells (mESC) to differentiate into the mesodermal cardiac lineage. The original EST consists of three different end-points (8): induction of cytotoxicity in 1) 3T3 cells (as a reference system for the maternal organism) and 2) mESC after 10 days of compound treatment determined by an MTT test, and 3) a differentiation assay, in which the inhibition of mESC differentiation into cardiomyocytes after 10 days of compound treatment is measured. For each of the three endpoints a 50% inhibition concentration is determined (IC50 for cytotoxicity and an ID50 for concentration of 50% inhibition of differentiation). The specific developmental element in this test strategy was the differentiation assay, in which the beating of cardiomyocytes was scored. Briefly, in the differentiation method, “hanging drops” containing 750 mESC each are cultured for 3 days in the presence of a concentration range of the test compound. During the hanging drop period, mESC form aggregates, called embryoid bodies (EB). At day 3 the EB are transferred to bacterial Petri dishes growing further in suspension medium containing the compound. On day 5, each EB is plated into one well of a 24 well plate where it attaches to the well surface and develops beating cardiomyocyte foci. Each plate is used to test a different compound concentration or control and the number of wells containing EB with beating cardiomyocytes is scored (figure 1). To optimize the test protocol a pre-validation study was performed in which ten compounds with different potencies of inducing developmental toxicity were tested by two laboratories (8). In this study, the prediction model described earlier by Spielmann *et al.* misclassified four out of ten compounds (7). An improved statistical prediction model was then introduced, which resulted in an accuracy of 93% for the same test compounds (8).
2.2. ECVAM validation of the EST

Following the prevalidation study, a larger validation study, funded by the European Centre for Validation of Alternative Methods (ECVAM), was performed following the ECVAM approach (8-10). This validation study was set up as an interlaboratory blind trial in which four laboratories of different international institutes tested 20 compounds, each with a different in vivo developmental toxicity potency, without the need of metabolisation (7 non-embryotoxic, 6 weakly embryotoxic and 7 strongly embryotoxic). The improved prediction model predicted the embryotoxicity of the compounds with an overall accuracy of 78%. For the strong embryotoxic compounds, the prediction was very good (100%) with a precision of 86%. Prediction is here defined as the likelihood that a positive prediction in EST correctly identifies a strongly embryotoxic chemical. Precision is defined as the proportion of chemicals which are strongly embryotoxic in vivo and correctly classified as such in the EST. The only strong embryotoxic compound not correctly predicted by the prediction model was methylmercury, which was classified as non-embryotoxic in 4 out of 8 experiments. Predictivity of the weak- and non-embryotoxic compounds was much lower (69% and 73%, respectively) producing a relatively high rate of false positive classifications. The test system was therefore deemed good in determining strong embryotoxicants, but the limitation was in discriminating between weak and non-embryotoxicants (9). The first validation test was very promising for the implementation of the EST as a test system to screen for developmental toxicants, however, it was based on a limited set of compounds, and further characterization with additional compounds and optimization of the model on the basis of results was deemed necessary.

2.3. EST: Post validation

In 2003, ECVAM organized an expert workshop to determine the applicability domain of the EST (11). Furthermore, it was evaluated if the validated method could be used as a screening test for lead compounds in the pharmaceutical industry and to detect the developmental toxicity of compounds in the chemical industry under REACH. Due to a number of limitations in the test system, the EST was not deemed fit to be accepted in regulatory context, but could be used to add supportive information (11). Furthermore, it was recommended that, to optimize the predictability and increase knowledge on the applicability of the EST, a number of improvements should be added to the EST, including:

1) development of a metabolic activation system,
2) differentiation of other specific lineages (e.g. neural, osteoblast etc.),
3) the existing chemicals database should be expanded with known in vivo embryotoxicants,
4) additional prediction models should be developed,
5) quantitative endpoints should be established,
6) in vitro/ in vivo concentration correlations should be considered,
7) the stability of the test compounds should be assessed.

As a response to the recommendations in this workshop, a selection procedure was performed by ECVAM and international experts, resulting in a group of thirty-one compounds which were selected to be included in a second screening study. Thirteen of these compounds, strong, moderate, mild or non-embryotoxic in vivo, were tested by two laboratories in the European ReProTect Consortium (12). The analysis for these compounds demonstrated only two out of thirteen compounds were classified correctly, indicating the improved prediction model was not capable of predicting developmental toxicity for this set. In a workshop organized by ReProTect and ECVAM, the unexpected results were discussed, and recommendations to modify the test system were advised (12). These included modifications to the EST protocol, addition of a metabolic system and new differentiation endpoints to be studied using molecular markers.

At the final stage of the ReProTect project, a feasibility study was performed in which ten compounds with a well documented toxicological profile were tested blindly in fourteen test systems, each predicting a certain endpoint within the reproductive cycle (13). The EST was included in this study as a predictor for developmental toxicity. Combining these fourteen assays resulted in a correct classification of nine out of ten compounds, with the EST contributing to these results.

2.4. Defining the applicability domain

Identifying the purpose for which the EST can be applied, is essential to determine the place of the test system within a testing strategy. The term applicability domain was defined by ECVAM as: Definition of the chemical classes and/or ranges of test method endpoints for which the model makes reliable predictions (10). According to this definition, the predictive value of the method for different chemical classes should assist in determining its applicability domain. Furthermore, the biological mechanisms incorporated within the system need to be characterized, to define the scope and limitations of the assay.

2.5. The applicability domain of compound classes for the EST

The first validation study of the EST proved it to be a valuable test system for predicting strong developmental toxicants, but less so in classifying weak- and non-developmental toxicants. Since then, many studies have been published by different research groups, describing the predictability of the test system for single compounds or compound classes containing compounds with different in vivo potencies for developmental toxicity.

Given the misclassification of the strong developmental toxicant methylmercury, the effects of other developmental toxic heavy metals, cadmium and arsenic (and metabolites), lithium, trivalent chromium and
hexavalent chromium were assessed in the EST (13-16). Cadmium, was misclassified as non-embryotoxic by the prediction model, due to higher cytotoxicity in the 3T3 cells or the absence of inhibition of beating cardiac foci (14). However, in the ReProTect feasibility study, cadmium was correctly classified as embryotoxic (13). Arsinite and arsenate compounds (including metabolites) were stated to be misclassified as non-developmental toxic in the EST, but multiple expert evaluations, including one on the selection of test chemicals for validation of in vitro embryo toxicity tests, concluded that inorganic arsenic is not a developmental toxicant in humans, supporting the classification observed in EST (14, 17, 18). Another study showed a correct classification for the heavy metals lithium and hexavalent chromium (strong embryotoxic), while confirming the misclassification of methylmercury (15, 16). In total, six of seven heavy metals were classified correctly with the EST.

The first study to evaluate the EST in a category approach using compounds within a distinct chemical class extensively studied the glycol ethers (19). Using the benchmark dose (BMD) approach the relative sensitivity of the EST within the class was compared to the relative developmental toxicity of the compounds within the class in vivo. The category approach assumes that when in vitro ranking of a class of compounds corresponds with the in vivo ranking, the test system will give reliable results for new compounds within the chemical class tested (19). It was shown that in the EST, all six glycol ether alkxy acid metabolites tested induced a concentration-dependent inhibition of cardiomyocyte differentiation. The ranking within the EST corresponded to the relative difference between the potencies, which were more pronounced in vivo, possibly due to differences in kinetics between in vitro and in vivo. This indicated the EST can be used as a tool to predict relative potencies of glycol ethers as a chemical class.

A second chemical class evaluated was valproic acid (VPA) and its analogs (20-22). In one study, VPA and five analogs with different potency were tested as a chemical class in the EST. Potency ranking of available in vivo data correlated well with potency ranking in the EST (21). Other groups studied VPA and six analogs using the original prediction model and the ReProGlo model, obtaining comparable results (20, 22).

These studies, using a category approach, show that the EST can be valuable in determining developmental toxicity within distinct chemical classes.

2.6. Implementation of the EST within testing strategies

Risk assessment determines the risks of human exposure to a compound or pharmaceutical. Regulatory risk assessment is currently based on animal toxicity studies. Intelligent testing strategies (ITS) are used to assess toxicological profiles, using decision schemes to determine the requirements of animal studies needed for hazard assessment (23). ITS use a stepwise approach (battery or tiered) with tests of increasing complexity to gather information on compound toxicity, weighing the information to determine further testing after each test performed, with the aim of limiting the number of animals needed. To further reduce the number of animals and decrease the time consuming process of toxicity testing, enhancing the use of in vitro developmental toxicity testing methods within ITS is desirable. In an ITS, such as the OECD conceptual framework, initial prioritization based upon all existing information is followed by non-testing information such as read-across and (Quantitative) Structure-Activity Relationships (QSAR) (24). Simple in vitro screening assays, such as receptor binding assays, will provide more mechanistic data and as a next step, more complex (multi)cellular in vitro models such as the EST can be studied. In this phase, multiple test systems each describing a part of development could be combined to give a more complete indication of possible untoward effects on embryogenesis, functioning as a prioritizing tool for determining whether in vivo developmental toxicity studies are necessary. The optimal place of EST within an ITS is dependent on a clear definition of its applicability domain, in terms of the biological processes represented in the assay.

3. QUANTITATIVE ENDPOINTS FOR THE EST

3.1. Alternative EST endpoints

Although the EST is a promising test system to determine developmental toxicity, it has its technical limitations. The test is laborious, taking 10 days of culture and scoring of the cardiomyocytes is subjective and time consuming. Shortening of test duration would increase the throughput of the test system. Furthermore, the EST evaluates effects on cardiomyocyte differentiation, but was intended to give an indication for any effects related to developmental toxicity. It has been shown that on day 10 of the culture, the percentage of cardiomyocyte foci derived from a single EB ranged between only 6-17% of the total amount of cells in the EB (25). To obtain more knowledge on the mechanisms of developmental toxicity, which can be evaluated with this test system, it would be preferable to study the toxic effects on all cell types within the system. In recent years, many proposals have been published to optimize, shorten and enhance the EST (12, 14, 26, 27). In addition, novel molecular endpoints have been proposed (20-22, 25, 28-35).

3.2. Optimization of the classical EST

During the ECVAM/ReProTect workshop many potential modifications to the existing method were proposed by industry (12). Modifications proposed were (among others), changes in media and time points, changes in the cytotoxicity test protocols or removal of the 3T3 cells in the testing strategy, changes in prediction model and alternative endpoints, such as molecular endpoints (FACS, gene expression models, alphaMHC-GFP-tagged cells) or a modification towards a two-class classification (embryotoxic or non-embryotoxic) (12). Apart from to these proposed modifications, many groups published their own optimized protocols.
Innovative approaches in the embryonic stem cell test

In an attempt to optimize the EST and make the method more high-throughput, Stummann et al. (36) developed a software program for digital movie analysis of contracting cardiomyocyte cultures. With this program, the beating frequency and areas of contracting cardiomyocytes can be measured more objectively and with more subtle endpoints compared to the subjective quantal scoring in the classical EST.

Van Dartel et al. (27) introduced a method to discriminate between compound effects on proliferation and differentiation. By exposing the cells in EST either from day 0 or from day 3 onwards, a difference in potency was found for the cytostatic agents 5-fluorouracil (5FU) and BrdU, but not for the developmental toxicants monobutyl phthalate (MBP) and 6-aminonicotinamide (6-AN). Further research showed that the cytostatic agents specifically affected cell proliferation during the first three days of the EST protocol, whereas the developmental toxicants MBP and 6-AN did not have a significant effect on proliferation. To discriminate further between proliferation and differentiation effects of compounds in the EST, van Dartel et al. proposed to perform the cytotoxicity assessment in EB at day 3 of the EST after day 0-3 exposure (replacing the 3T3 and ES cytotoxicity tests) and cardiac muscle foci counts after exposure from day 3-10 in the EST.

3.3. Molecular endpoints in the EST

To improve the highly subjective morphological scoring of the validated EST, multiple groups have proposed high throughput screening modifications to improve the EST by objectively and quantitatively determining molecular endpoints in the system for predicting developmental toxicants.

The first group to describe molecular endpoints in the EST, used fluorescence-activated cell sorting (FACS) analysis to study the quantitative expression of alpha-actinin and sarcomeric myosin heavy chain (MYC) (25). These markers were maximally expressed at day 7, allowing analysis at an earlier time point compared to the classical EST. In total, ten compounds were tested using the FACS-EST and results were compared to the validated EST, showing the comparable sensitivity in both systems (29). This improvement of the EST provides a more objective endpoint compared to morphological scoring, while on the other hand FACS analysis is time consuming and adds significant extra costs to performing the EST.

Another molecular endpoint used to predict developmental toxicity is gene expression (21, 37) Pellizer et al. were the first to use semi-quantitative reverse transcriptase (RT)-PCR to study key genes involved in cardiomyocytes development (Oct4, T, Nkx2.5 and alphaMHC) (37). Testing of retinoic acid (RA) and lithium chloride in this system showed a difference in mode of toxic action. Whereas RA had an effect on all four genes, lithium chloride only affected Nkx2.5 and alphaMHC genes which are expressed only later during differentiation. This was one of the first studies showing differences in mode of action for compounds in a mechanistic manner. De Jong et al. used Nkx2.5 and alphaMHC to compare gene expression after exposure with VPA and analogues at day 10 of the EST (21). In spite of the high variation in this experiment, a trend of inhibition for both markers was observed for VPA and its most potent analogues. These two studies suggest that measuring of single genes can give a complementary indication of mechanism of action compared to the classical endpoint of the EST.

Uibel et al. developed a completely new model, the ReProGlo assay, to predict early developmental toxicity using stem cells after 3 days of culture (20). In this model, the effects of compounds are measured on the canonical Wnt/β-catenin pathway, which is involved in early embryonic development. Cell viability and luciferase reporter activity is measured in a high-throughput system, where undifferentiated mESC are exposed to compounds for 24 hours. In total, 17 test chemicals with well characterized developmental toxic potential in vivo were tested in the ReProGlo system, including VPA and analogues and exposure to metabolites of cyclophosphamide derived from metabolism by primary mouse hepatocytes. All tested negative compounds were found negative and ten out of fourteen in vivo developmental toxicants were detected by the ReProGlo assay. This test system can be used to study the effect of compounds on the Wnt-signaling pathway, however, when a compound has a mechanism of toxic action not involving the Wnt-signaling pathway, it will not be detected.

3.4. Omics technologies as endpoints for the EST

Developmental toxic effects of compounds may not be detected using the EST, when cardiomyocyte differentiation is not specifically affected. At a molecular level, however, effects may be detected indicating a much broader range of effects. By using transcriptomics and proteomics, the mechanisms in the EST can be described more thoroughly, and the applicability of compound classes in the system can be defined more accurately. Van Dartel et al. extensively studied gene expression modulating effects of compounds on the EST using transcriptomics (30-34). In a first study, early gene expression changes (day 3-4 of the classical EST) accompanying cardiac differentiation were described, resulting in a set of 43 regulated genes describing cardiac differentiation (‘van Dartel gene set’) (33). The compound induced perturbed expression of this gene set was first studied for one model compound, monobutyl phthalate (MBP) at a concentration leading to an ID50 score in the classical EST, showing upregulation of gene sets related to pluripotency, proliferation and non-mesodermal differentiation and a downregulation of the 43 genes associated with cardiomyocyte differentiation. Studying the effects of MBP on the proteome, showed that MBP affected the expression of 33 proteins including downregulating the muscle protein myosin heavy chain (Myh) and affecting other development related proteins (38). A subsequent transcriptomics study focused on a longer exposure duration to MBP and 6-aminonicotinamide (6-AN), from day 3 onwards at 24 and 96h (34). Furthermore, differentiation without exposure was described over time from day 3 onwards after 0, 24, 48, 72 and 96h. In total, 1366 genes were significantly regulated.
over time and Principle Component Analysis (PCA) showed that the control samples were ranked in chronological order of culture duration, describing a so-called 'differentiation track' (34). MBP and 5-AN treated cultures were found to deviate significantly from this track, confirming the effect of these compounds on differentiation. To further validate this protocol, six and twelve compounds respectively with different developmental toxic potencies in vivo were tested in separate studies (at their ID50 concentrations in the classical EST), and the deviation from the differentiation track (day 3/4/5) after 24h exposure was studied (30, 32). Using the 'van Dartel gene set', 83% of the compounds were successfully predicted for developmental toxicity, indicating that the application of transcriptomics can improve the use of the EST for prediction of developmental toxic compounds. To further determine the performance of the EST, the differentiation track was used to study the developmental toxic effects of two compound classes: phthalates (n=3) and triazoles (n=3), each class containing compounds with different developmental toxic potencies (31). Differences between the two classes could be identified at the single gene level, and using the differentiation track the two classes could be discriminated. These results showed that using transcriptomics, compound classes could be identified and separated, giving information on mechanism of toxic action of a compound class and the applicability domain of the EST. To investigate the influence of the tested concentration, which is important for risk assessment, the triazole flusilazole was studied in a dose-response study (31). It was shown that developmental-related gene sets were affected at a lower concentration compared to cell-division related gene sets. Furthermore, upregulation of genes related to the antifungal mode of action of flusilazole could be detected, although at a higher concentration than the development related gene sets. In conclusion, these studies suggest that transcriptomic evaluation are informative for defining predictability and support the characterization of the applicability domain, improving the strength of the EST for regulatory testing strategies.

4. ALTERNATIVE DIFFERENTIATION MODELS

In EST only cardiomyocyte differentiation is assessed as an endpoint, and the test outcome is therefore uncertain as to the prediction of developmental toxicity of compounds with a mode of toxic action targeting other lineages of differentiation, such as neural, osteoblast or hepatocyte differentiation. An example of this is methylmercury, which was ranked as non-embryotoxic in the classical EST (9). In vivo, methylmercury is primarily a neurodevelopmental toxicant (39). Therefore it is plausible that a neural EST could detect the developmental toxic effects of this compound. Furthermore, many compounds have an effect on skeletal development in vivo, making an EST test system studying osteoblast or cartilage differentiation desirable. During the first ECVAM workshop it was advised to develop differentiation models studying specific lineages which could be used complementary to the EST in a testing strategy. Over the years, a range of protocols have been described to test these other specific lineages (11).

4.1. Neural differentiation EST

In vivo development of the nervous system is an intricate dynamic process, requiring coordinated expression of a range of complicated cellular and molecular events in a temporal- and region dependent process (40, 41). Cell types and key processes important in neural development are proliferation, differentiation of precursor cells into neurons and glia cells, migration, growth of axons and dendrites, synapse formation, myelination and programmed cell death (42). To incorporate all these endpoints in one in vitro system to study neurodevelopmental toxicity is highly complicated; therefore a battery of test systems, each describing a sub process in neural development is needed. This complexity has resulted in a range of diverse test systems published, studying compound perturbation on different aspects of neural development. The main focus currently rests on morphological endpoints, such as effects on axonal outgrowth and determining molecular endpoints predicting perturbation of neural differentiation.

Due to the misclassification of methylmercury in the classical EST, ECVAM tested this compound in a 25 day long neural differentiation version of the EST, using serum deprivation to induce neural differentiation (16). Cells were exposed to 100 nM methylmercury from day 0-14 and gene expression of eight well-known neural differentiation markers was studied by RT-PCR on days 0, 4, 9 and 14. The only significant change in gene expression was observed at day 14 for Mtap2. None of the other 7 markers were found to be significantly different compared to controls, indicating it is important to use an array of genes to study the effects in this kind of model to be able to study neural developmental effects.

Another neural differentiation model, quite similar to the classical EST was developed by Theunissen et al. (32). In this model, neural differentiation was induced by induction with a physiological concentration of retinoic acid and serum deprivation. It was shown that a 21 day neural differentiation protocol, a combination of two protocols earlier described by Okabe et al. and Bibel et al., could be reduced to a 13 day protocol, with only minor differences in morphology and expression of molecular markers at the end stage of the protocols(32, 43, 44). As a model to detect neurodevelopmental toxicity, a morphological screening assay was developed scoring the amount of neural outgrowth around embryoid bodies in combination with FACS analysis of three markers for pluripotency, early neural differentiation and mature neurons. In this test system, a non-cytotoxic concentration of 2.5 nM methylmercury significantly reduced the neural outgrowth FACS analysis only showed a significant effect on the expression of the early neural differentiation marker nestin. In a subsequent study, transcriptomics was used to study the effects of methylmercury on neural differentiation over time (45). Methylmercury was found to downregulate very early differentiation and pluripotency related processes and upregulate neural development related processes, as was observed earlier in vivo and in whole embryo culture transcriptomics studies studying methylmercury effects on early development (46, 47).
Kuegler et al. investigated critical processes mediating mESC differentiation into cells of the nervous system (48). A list of in total 197 mRNA markers was compiled from literature defining undifferentiated mESC, neural stem cells, astrocytes and the pattern of different neuronal and non-neuronal cells generated. These mRNA markers could be supportive in determining neurodevelopmental toxicity in neural differentiation cultures.

Due to the complexity of neural differentiation, it is advisable to first characterize the culture systems and then study the mechanisms behind neurodevelopmental toxicity in these models in order to define the applicability domain of these neural differentiation assays.

4.2. Osteoblast differentiation EST

Many developmental toxic compounds have an effect on skeletal development or bone formation and skeletal development is assessed in standardized regulatory developmental toxicity studies (49). An EST for osteoblast differentiation may contribute as an additional supplement to the battery of EST tests to determine developmental toxicity.

Zur Nieden et al. described a 30 day long method to differentiate mESC into mineralized osteoblasts using ascorbic acid, β-glycerophosphate and vitamin D3. From the second week of culture onwards, different osteoblast precursor genes were regulated, and mature bone specific genes were identified from the fourth week of culture (50). In a first study to determine developmental toxicity with this model, the effects of strong embryotoxicants all-trans retinoic acid and 5FU and non-embryotoxicant penicillin G were assessed by using multiple endpoints, of which two were found to give reliable results, namely: 1) morphometric IMAGE analyses of the mineralized osteoblasts expressing a black coloration under light microscopy and 2) calcium quantification in culture. Furthermore, it was shown that the differentiation protocol could be shortened to 14 days by using a monolayer culture, providing comparable results for the same three compounds (51). In another study by the same group the same model for evaluating developmental osteotoxicity was applied to identify the potential harmful effect of chloride on calcification (52).

4.3. Metabolism in the EST

One of the general problems of in vitro test systems is the lack of potential for metabolism within the test systems (53). Many known developmental toxicants are only toxic after biotransformation. In the EST, it was shown that methoxyethanol, a compound requiring metabolization into methoxyacetic acid in vivo to obtain embryotoxic properties, did not induce toxicity in the EST, while direct exposure to the metabolite methoxyacetic acid was toxic (21). The lack of metabolic capacity of ESC was recognized, and incorporating metabolic capacity to the EST system was marked as one of the possible improvements for the EST (11). Addition of S9 mix to the culture was unsuccessful, due to its cytotoxic effect after prolonged exposure in the EST (12). ECVAM introduced compound biotransformation in EST by using a genetically engineered mammalian cell line V79, transfected with CYP2B1 combined with the EST (54). The compound cyclophosphamide (CPA) is transformed by CYP2B1 into developmental toxic metabolites. Pre-incubation of CPA with these transfected cells for 24h and substituting the medium to differentiating cardiomyocytes made it possible to detect CPA as a developmental toxicant. Pre-incubating compounds for six hours in a murine or human primary hepatocyte culture, and subsequently exposing the EST differentiation culture to metabolites in this medium proved to be more successful (55). In this study, the compounds CPA and valpromide (VPM) were preincubated with primary hepatocytes. Exposure to pre-incubated CPA exposure medium resulted in 70 times lower ID50 values compared to non-activated CPA. Furthermore, a species difference was observed for bioactivation of VPD. Murine hepatocytes hardly formed any teratogenic VPD metabolites, however, human hepatocytes gave a significant conversion to the teratogen valproic acid.

In order to conduct proper risk assessment, apart from metabolism other kinetic properties of chemicals should also be taken into account, such as absorption, distribution and elimination of a compound. Especially, if the EST is to be applied to reliably extrapolate in vitro concentrations to in vivo exposures. Verweij et al. (56) were the first to show that physiological based kinetic (PBK) modeling could be a valuable tool to predict in vivo developmental toxic effect doses from in vitro effect concentrations obtained in the EST (56). In this study in vivo effect levels of 4 out of 5 compounds were predicted correctly in comparison with their in vivo lowest observed adverse effect level (LOAEL). More recent studies show that the concentration-response curves of a series of glycol ether alkoxyl metabolites in the EST can be extrapolated to in vivo dose-response curves which were in concordance with the embryotoxic dose levels measured in reported in vivo rat studies (21, 57). The first steps to introduce metabolism into the EST have been made, but further research and validation remains necessary.

4.4. Human EST

During the ECVAM workshop it was suggested that to further improve the applicability and safety assessment of chemicals by the EST, a major improvement could be obtained by the use human embryonic stem cells (hESC) in an EST model (11, 12). A number of differences in culture requirements, growth factors needed for differentiation and gene expression differences between mESC and hESC have been described over the years. For instance, mESC can be maintained in a pluripotent state by providing leukemia inhibitory factor (LIF), whereas hESC are not susceptible to this molecule and need fibroblast feeder layers to maintain pluripotency (58). Furthermore, it is known that within mESC and hESC, certain gene expression markers for pluripotency are common, such as Oct3/4 and Sox2, but distinct markers are only present in hESC, (SSEA-4, TRA-1-81 or in mESC (SSEA1) (58).

In recent years, a number of groups have studied the potential of hESC in compound developmental toxicity
assessment. The ECVAM group described the use of hESC and human fibroblasts in a cytotoxicity based system, comparable to the cytotoxicity tests in the classical EST (59). 5-FU and RA were tested in these systems and results were compared to historical classical EST data, showing comparable results. Cardiac differentiation of hESC was optimized and described by gene expression of pluripotency, early- and late cardiomyocyte markers. Later, the effects of 5-FU and RA on cardiomyocyte differentiation of hESC were evaluated, finding a concentration-dependent inhibition of the cardiomyocyte differentiation markers brachyury and GATA-4 (60). Another group studied the effect of six embryotoxic chemicals on the expression of multiple markers for ecto-, meso- and endodermal differentiation in hESC. The compounds studied displayed lineage specific effects on differentiation and showed that hESC may be a useful tool in compound developmental toxicity assessment (61).

Effects of developmental toxicants on neural differentiation of hESC have recently been described in a few studies. One study evaluated the effect of methylmercury on neural differentiation of hESC over 12 days and observed that exposure to 25 nM methylmercury caused a reduction in the gene expression of several neural differentiation markers, including MAP2, NCAM1 and NEUROD1 (62). In a previous study performed by the same authors using mESC, exposure to methylmercury only inhibited the gene expression of MAP2 at a concentration of 100nM but no other neural differentiation markers (16). This suggests the presence of interspecies differences in the mechanisms of action for developmental toxicity between hESC and mESC and provides support for developing hEST for possibly improved prediction of human adverse developmental effects.

5. FUTURE DEVELOPMENTS

Since its first definition, the EST has been evolving constantly towards a more perfected in vitro test system to detect developmental toxicity. Culture conditions within the test system have been improved and the applicability domain of the EST is being determined. In addition, multiple additional endpoints to the original prediction model have been studied and proposed, including molecular endpoints such as FACS analysis, RT-PCR and omics technologies. For use in chemical and pharmaceutical industry, further mechanization of the EST is required to increase high-throughput and the applicability domain of the EST should be further investigated. Furthermore, EST systems using differentiation into neural and osteoblast lineages as well as the use of hESC for developmental toxicity testing are being developed. Omics technologies could be used to further characterize the mechanisms of toxic action in both the classical EST as well as EST systems studying additional lineages and hESC differentiation. To be able to implement EST with novel molecular endpoints in developmental toxicity testing strategies for risk assessment purposes, further evaluation of predictability and applicability domains should be identified. In addition, the adversity of gene expression responses and their specificity as regards developmental versus maternal toxicity need to be addressed. The current and future efforts to improve the test systems will increase the value of embryonic stem cells used as well-characterized in vitro testing systems to predict developmental toxicity.

6. REFERENCES


2. P. C. Barrow: Reproductive toxicity testing for pharmaceuticals under ICH. Reprod Toxicol, 28(2), 172-9 (2009)


10. T. Hartung, S. Bremer, S. Casati, S. Coecke, R. Corvi, S. Fortaner, L. Gribaldo, M. Halder, S. Hoffmann,
Innovative approaches in the embryonic stem cell test


Innovative approaches in the embryonic stem cell test


Innovative approaches in the embryonic stem cell test


**Key Words:** Embryonic stem cell test (EST), Alternative methods, Toxicology, Review

**Send correspondence to:** Aldert H. Piersma, National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, P.O.Box 1.3720 BA Bilthoven, The Netherlands, Tel: 31 30 274 2526, Fax: 31 30 274 4446, E-mail: Aldert.Piersma@rivm.nl

http://www.bioscience.org/current/vol17.htm