Toxicology Letters 230 (2014) 356–367

Contents lists available at ScienceDirect

Toxicology Letters

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

An integrated approach for detecting embryotoxicity and developmental toxicity of environmental contaminants using *in vitro* alternative methods

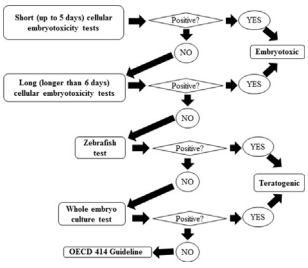
Miguel A. Sogorb[a,*], David Pamies[a,b], Joaquín de Lapuente[c], Carmen Estevan[a,d], Jorge Estévez[a], Eugenio Vilanova[a]

[a] *Unidad de Toxicología y Seguridad Química, Instituto de Bioingeniería, Universidad Miguel Hernández de Elche, Elche, Spain*
[b] *Johns Hopkins University, Bloomberg School of Public Health, CAAT, Baltimore, USA*
[c] *Centre de Recerca en Toxicologia (UTOX-CERETOX), Parc Científic de Barcelona, Barcelona, Spain*
[d] *Chemical Assessment and Testing Unit, Institute for Health and Consumer Protection, DG Joint Research Centre, European Commission, Spain*

HIGHLIGHTS

- Stem cells under differentiation are useful model for testing embryotoxicity *in vitro*.
- Zebra fish and WEC are widely used for testing teratogenicity *in vitro*.
- Omic approaches contributed to enhance predictivity of developmental toxicity.
- A single test would not show enough predictivity for screening developmental toxicity.
- A tiered approach strategy would reduce bioethical concerns in developmental toxicity.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 29 September 2013
Received in revised form
10 December 2013
Accepted 27 January 2014
Available online 3 February 2014

Keywords:
Embryotoxicity
Teratogenicity
Alternative tests
Embryonic stem cell test
Zebrafish test
WEC

ABSTRACT

The main available alternatives for testing embryotoxicity are cellular tests with stem cells and *in vitro-ex vivo* tests with embryos. In cellular tests, the most developed alternative is the embryonic stem cell test, while the most developed tests involving embryos are the zebrafish and whole embryo culture test. They are technically more complex than cellular tests, but offer the advantage of determining the expectable phenotypic alteration caused by the exposure. Many efforts are currently being made, basically through proteomic and genomic approaches, in order to obtain improvements in predictivity of these tests. Development is a very complex process, and it is highly unlikely that a single alternative test can yield satisfactory performance with all types of chemicals. We propose a step-wise approach where model complexity, and consequently technical skills and economical costs, gradually increase if needed. The first level would be run short cellular assays to detect effects in early differentiation stages. The second level would involve longer cellular embryotoxicity tests to search embryotoxicants that have an effect on late differentiation stages. The third stage would consider tests with embryos because they allow the determination of hazards based on molecular and morphological alterations, and not only on differentiating cells.

* Corresponding author. Tel.: +34 966658506; fax: +34 966658511.
E-mail address: msogorb@umh.es (M.A. Sogorb).

0378-4274/$ – see front matter © 2014 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.toxlet.2014.01.037

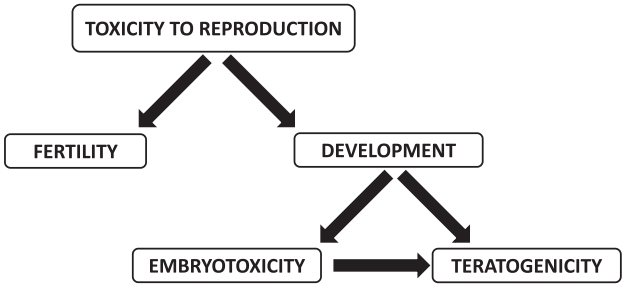


Fig. 1. Relationships among toxicity to reproduction, developmental toxicity, embryotoxicity and teratogenicity.

1. Introduction

According to their definitions, developmental toxicity is adverse effects on the growing organism from the embryonal state to the time of an individual's sexual maturation, embryotoxicity is the toxic effects in the progeny between conception and the foetal stage, while teratogenicity is structural malformations or defects in offspring after the embryogenesis period. Thus, embryotoxicity and teratogenicity are both considered to be toxic effects on development. Toxicity to reproduction is a term that encompasses alterations in fertility and development caused by chemicals. Fig. 1 outlines the relationships among these different toxicity classes included under the term "toxicity to reproduction". The developmental toxicity of environmental contaminants is an issue that causes much concern for society, and toxicologists must possess safe and reliable procedures to test it. The main protocols normalised by the Organization for Economic Co-operation and Development (OECD)[1] to test developmental toxicity include the following guidelines: 414, for testing teratogenicity; 416, for testing toxicity to reproduction (fertility plus development) in two generations; 421, a screening study for testing developmental toxicity; 422, for testing toxicity to reproduction simultaneously with repeated dose toxicity; 426, for testing neurodevelopmental toxicity (Estevan et al., 2011). OCDE guideline 443, for testing reproductive toxicity in one-extended generation, was added to the list in 2012.

The main inconveniences of these OECD guidelines are that they are expensive, time-consuming and use a large number of animals, plus the associated bioethical and social concerns. Table 1 shows the economic cost and the minimum number of animals requested for applying OECD guidelines to test toxicity to reproduction. Indeed, the estimated economic cost ranges between €54,600 for guideline 421 and €1,100,000 for guideline 426 (Rovida and Hartung, 2009). Moreover, the number of animals required ranges between 412 for guideline 422 and 3,200 for guideline 416 (Rovida and Hartung, 2009). It is remarkable that testing embryotoxicity alone is not considered among the various OECD guidelines. This part of developmental toxicity must be assayed necessarily in the long general test of toxicity to reproduction, where fertility, embryotoxicity, teratogenicity and development are assessed in the same test.

All these data suggest that cheap and reliable methods for testing developmental hazards of environmental contaminants and the subsequent risk assessments would be welcomed. In this scenario, alternative and *in vitro* methods for testing developmental toxicity might play a relevant role. The term alternative method is assigned to those methods used to study toxicity that Reduce, Refine or Replace (3Rs) animal experimentation (Russell and Burch, 1959).

[1] An integrated approach for detecting embryotoxicity and developmental toxicity of environmental contaminants using in vitro alternative methods.

Table 1
Economical cost and number of animals needed to apply the OECD Guidelines for testing reproductive toxicology (data taken from Rovida and Hartung, 2009).

OECD guideline	Purpose	Animals	Estimated cost (€)
414	Teratogenicity	784	63,100 (rats) 92,500 (rabbits)
416	Reproductive toxicity in two generations	3,200[a]	328,000
421	Screening test for reproductive and developmental toxicity	560	54,600
422	Combined repeated dose toxicity study with the reproduction/developmental toxicity screening test	412	92,000
426	Neurodevelopmental toxicity study	1,400[a]	1,100

[a] Considering all the discarded pups.

The OECD guidelines for testing chemicals include 20 different *in vitro* methods to test various aspects of toxicity, such as dermal and ocular impairments, genotoxicity and endocrine disruption. Yet none is devoted to toxicity to reproduction, developmental toxicity or embryotoxicity.

Nevertheless, several *in vitro* alternative methods have been developed for testing embryotoxicity and teratogenicity. They all offer high predictivity and concordance with *in vivo* test results, meet the 3Rs requirements, and are cheaper than *in vivo* tests. The main alternative tests for assessing embryotoxicity are based on the *in vitro* study of alterations in cellular differentiation, while alternative methods for testing teratogenicity are based on the *in vitro* or *ex vivo* exposure of whole embryos, with further analyses of the alterations caused by the chemical being assessed (Table 2). Whole Embryo Culture (WEC), micromass (MM) and Embryonic Stem cell Test (EST) have been validated in blind studies (Genschow et al., 2002) and a study is currently under way with zebrafish (Gustafson et al., 2012). These studies have been carried out under the typical perspective of verifying reproducibility and relevance. However, mechanistic validation has been recently proposed as a tool to proceed in order to generate valuable information for decision-making on the basis of *in vitro* results (Hartung et al., 2013a).

This review describes the main alternative *in vitro* tests available for determining embryotoxicity and teratogenicity of environmental contaminants and proposes an integrated approach with a step-wise strategy that would allow the assessment of developmental toxicity on the basis of these robust *in vitro*-alternative tests. The proposed integrated approach is outlined in Fig. 2. In this step-wise approach, model complexity and, consequently technical skills and economical costs, can be gradually increased if required. The first level would be to run short (up to 5 days) cellular assays to detect embryotoxicants that exert effects on early differentiation stages. The second level would entail longer (more than 6 days) cellular embryotoxicity tests to search for embryotoxicants that have effects on late differentiation stages. If positive effects are proven at either of these two levels, the environmental contaminant may be considered an embryotoxicant. If negative

Table 2
Main alternative methods available for assessing developmental toxicity.

Embryotoxicity	Teratogenicity
Embryonic Stem cell Test (EST)	Micromass test (MM)
Mouse embryonic stem cell adherent cell differentiation and cytotoxicity (ACDC) test	Whole embryo culture test (WEC)
Assays with human embryonic stem cells	Zebrafish embryonic development test
Rat cerebellar granule cells	

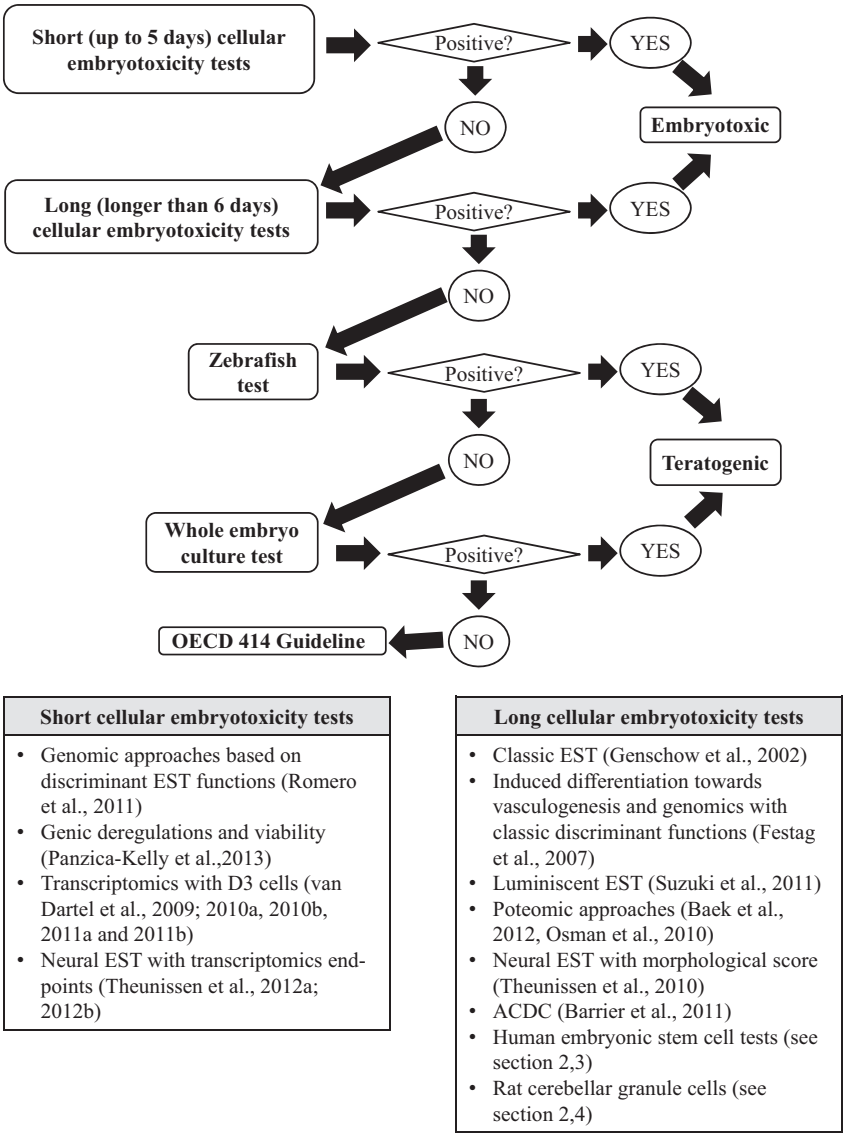


Fig. 2. General outline of an integrated tiered strategy of embryotoxicity and teratogenicity tests to screen developmental toxicants based on alternative methods. The flow diagram indicates the process that can be followed to identify the embryotoxic and teratogenic hazards of environmental contaminants.

results are obtained, then the zebrafish test is proposed as a third stage in the process because it would determine hazards based on molecular and morphological alterations in a whole embryo, and not only in differentiating cells. A second type of assays (indeed the fourth level in the scale) with more complex embryos (rat) would entail WEC assays. Finally, the last stage would implicate *in vivo* assays following OECD guideline 414. If positive results are obtained in any stage, the tested chemical should be labelled as embryotoxic or teratogenic, and any remaining assays should be avoided. If, however, negative results are observed in all the tests, the whole *in vivo* assay should be avoided depending on specific regulations or, alternatively, a focused and directed approach, e.g., performing a limit assay, might suffice. A battery of complementary tests (including cellular tests and test with embryos) has been successfully employed to correctly predict the developmental toxicity of eleven of twelve chemicals with different mechanisms of toxic action with only a false negative (Piersma et al., 2013).

2. *In vitro* assays for testing embryotoxicity

The main *in vitro* embryotoxicity tests are based on the exposure of stem cells under differentiation to a chemical, and the further assessment of alterations in the differentiation caused by the exposure. The assessment of such alterations can be made by different approaches; e.g., histological, genomic, proteomic, functional alterations, etc.

2.1. Embryonic Stem cell Test (EST)

The embryonic stem cell test (EST) is probably the most developed *in vitro* methodology for testing embryotoxicity. It passed a blind inter-laboratory validation test sponsored by the European Union Reference Laboratory for Alternatives to Animal Testing (EURL-ECVAM), which determined that, although the method was scientifically valid and validated, it was not ready to be used for regulatory purposes (ESAC, 2002).

The method is based on studying alterations in the differentiation of D3 mouse embryonic stem cells, which spontaneously generate contractile cardiomyocytes after 10 days of culture as an embryoid body (see a beating cardiomyocyte generated during an EST run at the end of the video in http://www.youtube.com/watch?v=-gPxGM606Ms). This alteration in D3 differentiation is complemented with other parameters, such as alterations in the

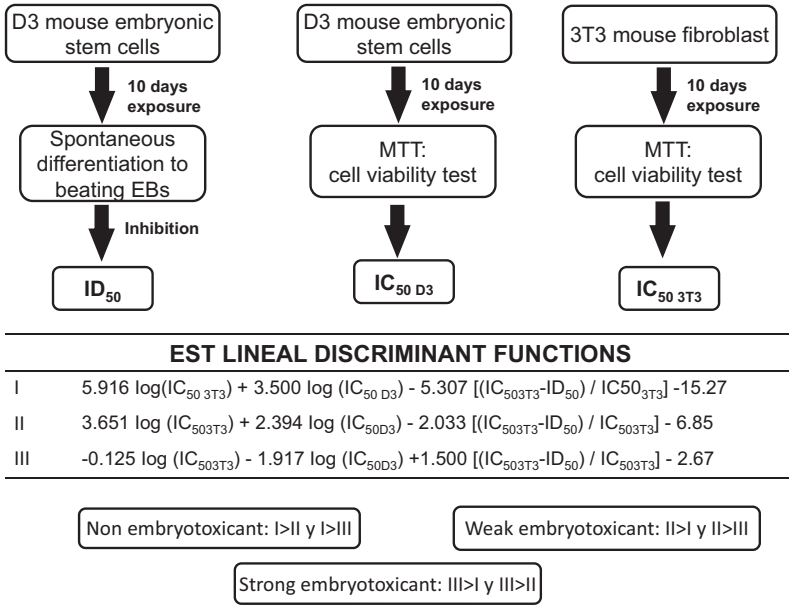


Fig. 3. General outline of the embryonic stem cell test. ID_{50} = concentration of the chemical causing a 50% reduction in the differentiation of undifferentiated D3 cells towards beating cardiomyocytes (see a beating cardiomyocyte at: http://www.youtube.com/watch?v=-gPxGM606Ms). IC_{50D3} and IC_{503T3} = concentrations of chemicals causing a 50% reduction in the viability of D3 and 3T3 cultures, respectively. MTT = method for testing cell viability based on the reduction of tetrazolium formazan. The linear discriminant functions and criteria to assign the embryotoxicity category were extracted from Genschow et al., 2002.

viability of both these D3 cells and 3T3 mouse fibroblasts. Thus, EST is based on the determination of three different end-points defined as: (1) ID_{50}, the concentration of chemical that causes a 50% reduction in the differentiation of undifferentiated D3 cells towards beating cardiomyocytes; (2) IC_{50D3}, the concentration of chemicals that cause a 50% reduction in the viability of D3 cultures; (3) IC_{503T3}, the concentration of chemicals that bring about a 50% reduction in the viability of 3T3 cultures. These three end-points are further analysed by three empiric discriminant lineal functions that allow the assessed chemical to be classified according to its *in vivo* embryotoxicity into one of the three following categories: (1) strong embryotoxicants; (2) weak embryotoxicants; (3) non-embryotoxicants (Genschow et al., 2002). Fig. 3 outlines the general EST methodology procedure. This classic validated EST procedure might be suitable to be included within the long cellular embryotoxicity tests proposed for the strategy for testing embryotoxicity outlined in Fig. 2.

The main inconvenience of EST is the lack of accuracy in the measures of alterations in D3 cells differentiation. Alterations in differentiation are based on the morphological determinations made by optical microscopy of the beating in generated cardiomyocytes. This quantification procedure contemplates only two different situations, the cardiomyocyte beats or the cardiomyocyte does not beat, but does not consider variables as the total area of beating, the frequency of beating, the morphology of the cardiomyocyte, among others. Thus, the quantification of alterations in differentiation is performed with uncertainty, which confers EST a broad scope for improvement, especially for discrimination between weak and non-embryo toxicants. Indeed, EURL-ECVAM has put forward several suggestions to enhance the predictivity of this methodology, which specially include the necessity of developing molecular approaches for the quantification of cell differentiation alterations (Spielmann et al., 2006).EST has proven its reliability for the screening of embryotoxicity of lots of environmental contaminants as, among others, environmental oestrogens contaminants (Kong et al., 2013), triazole fungicides (Piersma et al., 2013; de Jong et al., 2011a); phenols (Strikwold et al., 2012); phtalates (Piersma et al., 2013); glycol ethers (de Jong et al., 2009); pesticides as endosulfan and glufosinate (Piersma et al., 2013) and 2,3,7,8-tetrachlorodibenzo-para-dioxin (Neri et al., 2011). EST has been also used for the study of embryotoxicity of engineered nanomaterials (ENM) finding that cobalt ferrite nanoparticles coated with gold and silanes should be classified as non-embryotoxic, while others as gold salt, cobalt ferrite salt, cobalt ferrite salt coated with silanes and gold coated with coated with hyaluronic acid should be classified as weak embryotoxicants under EST criteria (Di Guglielmo et al., 2010).

2.2. The genomic approaches proposed to enhance EST performance

The expression of gene biomarkers of differentiation has been researched as a molecular end-point to study differentiation alterations in the EST. In line with this, by using D3 cells and the same prediction models as in a regular EST (Fig. 3), Romero et al. (2011) proved that it is possible to discriminate among the embryotoxicities of 5-fluorouracil (a strong embryotoxicant), 5,5-diphenylhydantoin (a weak embryotoxicant) and saccharin (a non-embryotoxicant) by employing the concentration that causes a 50% reduction in the expression of the following genes as ID_{50}: the patatin-like phospholipase domain containing 6 (*Pnpla6*), α-foetoprotein (*Afp*), histone deacetylase 7 (*Hdac7*), vascular endothelial growth factor A (*Vegfa*), foetal liver kinase 1 (*Flk1*) and nestin (*Nes*). The other significant improvements of this proposal included cutting exposure time to 5 days (half the duration of a regular EST), employing monolayer cultures, avoiding technically complex embryoid bodies, and monitoring differentiation alterations of cellular lineages other than the mesoderm, the main tissue of cardiomyocytes, which is another of the EURL-ECVAM's suggestions (Spielmann et al., 2006) to improve the classical EST. The total length of this procedure was five days and therefore this methodology is suitable to be included within the short cellular embryotoxicity test considered for the second stage of the strategy showed in Fig. 2.

Festag and co-workers (2007) also introduced disturbances of vasculogenesis and/or angiogenesis as end-points in the EST. They used the linear discrimination functions of the classic EST (see Fig. 3) by substituting the alterations in the generation of beating

cardiomyocytes for alterations in the expressions of the platelet-endothelial cell adhesion molecule-1 and vascular endothelial cadherin genes. This methodology was able to correctly predict the embryotoxicity of six different model compounds (two for each category) using 10 days of differentiation and therefore it might be considered for second stage of tests displayed in Fig. 2 (long cellular embryotoxicity tests).

The potency of inducing the embryotoxicities of six structurally related valproic acid analogues was determined by the classic EST approach and by the gene expression analysis of cardiac markers (NK2 homeobox 5 and myosin heavy chain). A sound agreement was found between the *in vivo* results and the *in vitro* results obtained when recording beating cardiomyocytes and using a genomic approach (de Jong et al., 2011b). These two genes were also used as end-points to establish the disruption of D3 cells differentiation into cardiomyocytes caused by exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin, which allowed to establish the potential role of dioxin as an embryotoxic contaminant (Wang et al., 2010). Genomic approaches also offer added value which might help our understanding of embryotoxicity mechanisms. In this way, Neri and coworkers (2011) reported that the abnormalities caused by 2,3,7,8-tetrachlorodibenzo-p-dioxin during heart development must be attributable to reductions in the ATP levels caused by alterations in mitochondria arrangements, inductions of the enzymes involved in phase I of biotransformation, and down-regulations in oxidative phosphorylation pathways. Lack of ATP also seems to be responsible for the misalignment or disarrangement of myofibrillar organisation (Neri et al., 2011).

Heart and neural crest derivatives expressed transcript 1 (*Hand1*) and cardiomyopathy associated 1 (*Cmya1*) are both genes that are greatly involved in heart development. A stable transgenic D3 mouse embryonic stem cell line was generated with the *Hand1* and *Cmya1* promoters upstream of the luciferase reporter gene. In this way, changes in the expressions of these genes will be coincident with luciferase activity during cardiomyocyte differentiation. The discriminant functions of a classic EST proved to be good predictive models when used to test the embryotoxicity of 24 different model chemicals by utilising alterations in both the viability of 3T3 and these transgenic D3 and the expression of *Hand1* and *Cmya1* as end-points for biomarkers of differentiation (Suzuki et al., 2011). This variation in the EST conferred the advantage of it being suitable for high throughput screening because the whole methodology was based on luminescent measures with 96-well plates, which cuts of exposure time to 6 days. However the main inconvenience was that this methodology only records alterations in cardiomyocyte differentiation, and it avoids other lineages as potential targets of embryotoxicity. This luminescent variant of the EST was further validated in a preliminary cross-laboratory study, which showed a good correspondence with all four laboratories involved (Suzuki et al., 2012) and deserves to be included within the methodologies for testing embryotoxicity with cellular procedures based on long exposures (Fig. 2).

Panzica-Kelly et al., 2013 developed a decision tree model to assess the embryotoxicity of chemicals based on two end-points obtained using only D3 cells (avoiding 3T3 cultures). These two end-points were alterations by 50% of cell viability (lower than 22 μM, between 22 and 200 μM, or higher than 200 μM) and if this concentration was able to cause a deregulation of at least 10% in at least 8 of the genes included in a set with 12 developmentally regulated gene targets. This decision tree enabled the correct prediction of the teratogenicities of 45 of the 65 chemicals included in a trial set. The main advantage of this methodology is that it reduces exposure time to 4 days (considered as short exposures procedures in Fig. 2) and avoids the use of a second cell type. However, it allows chemicals to be classified into only two categories, instead of three as determined by a classic EST.

Alterations in the gene expression profile of D3 cells under differentiation, in combination with alterations in viability and acetylcholinesterase inhibition of D3 and 3T3 cells, have also been used as a tool to predict potential embryotoxic hazards of organophosphorous insecticides. Accordingly, chlorpyrifos was able to strongly alter the expression of biomarkers genes of differentiation at concentrations on the limit to cause cholinergic toxicity to the mother (Estevan et al., 2013). The assessment of the situation with the chlorpyrifos metabolites, chlorpyrifos-oxon and 3,5,6-trichloro-2-pyridinol, allowed to conclude that, in exposure scenarios where low liver bioactivation was present (i.e., after inhalatory or dermal exposures) it is possible to reach chlorpyrifos concentrations that strongly alter the gene profile expression with low or even in absence of maternal toxicity. All these facts might explain the contradictory *in vivo* and epidemiological studies conducted into the capability of chlorpyrifos to induce developmental toxicity because most of the toxicity studies with animals are run using oral administration (with high liver bioactivation), while this is not the expectable exposure route for humans.

Transcriptomics is a promising tool for enhancing EST predictivity. Indeed, two different sets of genes, "EST biomarker genes" (26 genes) and "Van_Dartel_heartdiff_24 h" (43 genes), related with the differentiation process in D3 cells, were able to predict, by a principal component analysis (PCA) and enriched gene ontology, the embryotoxicity of the various model compounds employed in the validation study of EST, such as retinoic acid, penicillin-G, methoxyacetic acid, 6-aminonicotinamide and 5-fluorouracil, and other chemicals with well-known *in vivo* embryotoxicity; e.g., several triazol and phthalate analogues, carbamazepine, methylmercury, etc. (van Dartel et al., 2009, 2010a,b, 2011a,b). The main technical advantage of these procedures is that they use a single concentration of the tested chemical and cut exposure time to 24 h (short cellular test in Fig. 2), but the use of a single dose does not allow establishing a dose-response relationship.

2.3. Proteomic approaches proposed for enhancing EST performance

The main proteomic approach for monitoring differentiation of D3 towards cardiomyocytes is the record of proteins with majoritarian expression in cardiac muscle. A variant of the classic EST was developed whose end-point was to follow alterations in differentiation through changes in the expression of myosin heavy chain and actinin. This methodology yielded good concordance between *in vivo* embryotoxicity and *in vitro* predictions with 10 different model compounds (4 strong embryotoxicants, 3 weak embryotoxicants and 4 non-embryotoxicants) (Seiler et al., 2004; Buesen et al., 2009), and six closely related congeners of valproic acid (Riebeling et al., 2011). A similar approach was used to screen embryotoxic chemicals using the expression of *Tuj-1* (a well-known early neuronal marker) as an end-point to follow differentiation. This methodology allowed the classification of methyl mercury, valproic acid, sodium arsenate and sodium arsenite as embryotoxicants (Baek et al., 2012). Again, the main inconvenience of these procedures is that they are based in the assessment of alterations in differentiation to only one lineage instead to all the three main cellular lineages found in a differentiating embryoid body.

The analysis of the whole proteome of D3-differentiating cells revealed that their exposure to embryotoxicant monobutyl phthalate caused the up-regulation of 57 proteins and the down-regulation of 42 others. Most up-regulated proteins were correlated with cardiomyocyte functionality, while down-regulated proteins were principally pluripotency markers, chaperones and ribosomal proteins (Osman et al., 2010). These data showed a positive correlation with the expression of related genes at the transcriptomic level (Osman et al., 2010). All these data revealed the great potential

of proteomics as an end-point for testing embryotoxicity using D3 cells.

Both proteomic approaches would be included as long cellular embryotoxicity test (Fig. 2) since the total length of the exposure ranged between 7 (Baek et al., 2012) and 10 days (Osman et al., 2010). However, it seems that a similar strategy might be also suitable to be used with shorter embryotoxicity cellular tests.

2.4. The neural Embryonic Stem cell Test (nEST)

The spontaneous differentiation of D3 mouse embryoid bodies yields beating cardiomyocytes, where the predominant lineage is mesoderm. This is one of the limitations of the EST because it might not detect the embryotoxicants that exert their effects on other lineages. In order to overcome this problem and to enhance the capability of the EST to detect neurodevelopmental toxicants, a protocol based on the exposure of differentiating cells to distinct trophic factors and cells culture media was developed. It allowed the generation of a cell culture with a predominant neuroectodermal lineage and low proportions of endodermal and mesodermal differentiated cell types in 13 days (a similar time to the classic EST) (Theunissen et al., 2010). The end-points with this neural EST can be either morphological at the end of the exposure (long-term cellular embryotoxicity test (Fig. 2)) or transcriptomics after shorter exposures (short-term cellular embryotoxicity test (Fig. 2)). This differentiation process was seriously perturbed by exposure to methyl mercury, which suggests the possibility of it being a good model to screen neurodevelopmental embryotoxicants (Theunissen et al., 2010). A transcriptomic analysis with microarray and further PCA identified an optimised set of 29 genes, which conferred 84% prediction accuracy in a set of 10 different model compounds (Pennings et al., 2012). The combination of transcriptomics with gene ontology biological processes and morphological scoring (assessment of the extent of neurite outgrowth) allowed the successful screening of the neurodevelopmental embryotoxicity of different chemicals: e.g., acetaldehyde, carbamazepine, flusilazole, monoethylhexylphthalate, penicillin G, phenytoin, cyproconazole, hexaconazole and valproic acid, among others (Theunissen et al., 2012a,b). It has been recently suggested that a combined approach incorporating a regular EST and nEST may improve developmental toxicant detection in individual assays (Theunissen et al., 2013).

2.5. The mouse embryonic stem cell adherent cell differentiation and cytotoxicity (ACDC) assay

Other cellular assays are available that utilise only J1 cells, another pluripotent mouse embryonic stem cell. This methodology, called the ACDC assay, is based on the quantification of the myosin heavy chain protein as a marker of cardiomyocyte differentiation, which was further corrected for cell number, thereby separating cytotoxicity and effects on differentiation (Barrier et al., 2011). The total length of the test is 10 days and therefore must be considered as an assay for testing the embryotoxic effects caused after long exposures (Fig. 2). This system was able to initially discriminate the relative embryotoxic potency of acetic acid, 5-fluorouracil and bromochloroacetic (Barrier et al., 2011). This methodology further demonstrated its suitability for high throughput in the assessment of 309 chemicals, which allowed the initial characterization of metabolic and regulatory pathways by which some environmental chemicals may act to disrupt embryonic stem cell growth and differentiation (Chandler et al., 2011). The main advantage of this system is that it permits both determinations (differentiation and viability) in a single assay on the same platform. However, the end-point proposed for assessing differentiation is based on a biomarker of differentiation to a single lineage (myosin heavy chain as a biomarker of mesoderm-derived cardiomyocytes).

2.6. Assays with human embryonic stem cells

WA09 human embryonic stem cells differentiated *in vitro* in a 3D system with gene expression profiling suggested a predominantly forebrain-like development, and indicated that this model might be suitable for assessing neurodevelopmental alterations. Indeed, this gene expression profile was severely altered at non-cytotoxic concentrations of the known developmental neurotoxicant methyl mercury and by chemically inert polyethylene nanoparticles (Hoelting et al., 2013). Another promising 3D model for testing neurodevelopmental toxicity is human neural progenitor cells, which grow as neurospheres. Developmental neurotoxicants methylmercury chloride and mercury chloride shortened migration distance and reduced the number of neuronal-like cells in differentiated human neural progenitor cells (Moors et al., 2009). In addition, other 3D models generated from induced pluripotent stem cells also have been presented as suitable candidates to evaluate developmental neurotoxicity effects (Hogberg et al., 2013).

Human embryonic stem cells can be differentiated *in vitro* by mimicking the developing neural plate and neural tube, and the process can be recorded according to the profile expression of genes that are relevant to early neural development. Retinoic acid altered the morphology of the resulting structures with similar gene expression changes to those induced by retinoic acid *in vivo* (Colleoni et al., 2011), suggesting that this model might be relevant for screening neurodevelopmental chemicals which act in the early steps of the process.

It has been revealed that a model with a human neural stem cell line derived from umbilical cord blood can discriminate between developmental and non-developmental toxicants using cell viability, proliferation, apoptosis, and expression of cell type-specific markers as end-points after a 48 h exposure with a set of the following chemicals: sodium tellurite, methylmercury chloride, cadmium chloride, chlorpyrifos, and L-glutamate, acetaminophen, theophylline, and D-glutamate (Buzanska et al., 2009).

All these assays with human embryonic stem cells involve exposures longer than 6 days and therefore must be considered as long cellular embryotoxicity tests according to the step wise strategy described in Fig. 2.

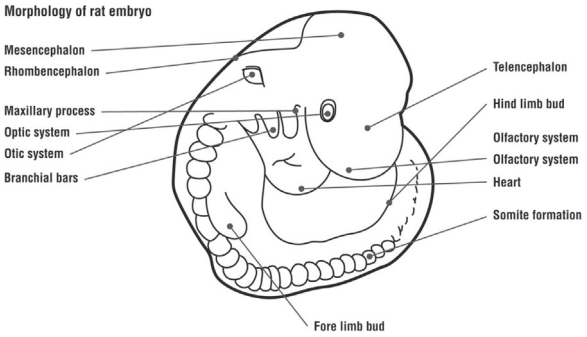
2.7. Rat cerebellar granule cells

Primary cultures of rat cerebellar granule cells are also a widely used model for testing developmental neurotoxicity. Neurodifferentiation alterations are usually followed by analysing changes in the gene expression profiles of neuron precursor, neuron and astrocytic markers, as proven with model chemicals as methyl mercury chloride, lead chloride, valproic acid and tri-methyl tin chloride (Hogberg et al., 2010). This same approach also demonstrated that paraquat parathion, dichlorvos, pentachlorophenol and cyclohexiimide can induce developmental neurotoxicity (Hogberg et al., 2009). The exposures needed to detect embryotoxicity using rat cerebellar granule cells are longer than 10 days and therefore the methodology must be included in the second step of the step-wise strategy described in Fig. 2.

3. *In vitro* assays for testing teratogenicity

3.1. Micromass test (MM)

The predictivity of this method falls in observations of alterations in differentiation towards cartilage of primary cultures of



MM LINEAL DISCRIMINANT FUNCTIONS	
I	6.65 x log (ID_{50}) – 9.49
II	6.16 x log (ID_{50}) – 8.29
III	-1.31 x log (ID_{50}) – 1.42

WEC LINEAL DISCRIMINANT FUNCTIONS	
I	15,37 log(IC_{max}) + 21 [(IC_{503T3}-$IC_{NOEC\ TMS}$) / IC_{503T3}] - 23,58
II	17,71 log(IC_{max}) + 27 [(IC_{503T3}-$IC_{NOEC\ TMS}$) / IC_{503T3}] – 32,37
III	4,21 log(IC_{max}) + 9,3 [(IC_{503T3}-$IC_{NOEC\ TMS}$) / IC_{503T3}] – 4,23

Non embryotoxicant: I>II y I>III

Weak embryotoxicant: II>I y II>III

Strong embryotoxicant: III>I y III>II

Fig. 4. Morphology of rat embryo, prediction models and criteria for the assignment of embryotoxicity in the micromass (MM) and Whole Embryo Culture (WEC) tests. IC_{503T3} = concentration that reduces viability of 3T3 cells to 50% after exposure according to the protocol; ID_{50} = concentration that reduces the differentiation of primary culture of limb bud cells to cartilage to 50% after exposure according to the MM protocols; $IC_{NOEC\ TMS}$ = the lowest concentration with no observed effect on the total morphological score (see Table 3); IC_{max} = The lowest concentration that causes the maximum malformation rate. The picture was taken from Pamies et al., 2011 (with permission). Linear discriminant functions and criteria for assigning the embryotoxicity category were extracted from Genschow et al., 2002.

fore limb buds of rat embryos by day 14 of gestation (Fig. 4) after a 5-day exposure. This differentiation is a critical step for the morphogenesis of the skeleton and for other processes such as cell proliferation and differentiation, cell-to-cell communication and cell-to-extracellular matrix interactions (Pamies et al., 2011). After having obtained the concentration of the chemical that capable of inhibiting 50% of differentiation (ID_{50}), lineal discriminant functions allow to classify the assessed chemical into the same three embryotoxicants categories as in the EST (Fig. 4). This MM test passed an inter-laboratory validation study and EURL-ECVAM concluded that the method was ready for detecting only strong embryotoxicants (ESAC, 2002).

Alternatively, alterations in the viability of these primary cultures can also be used as end-point for screening teratogenic chemicals. Certain studies, performed with fungicide carbendazim (Minta et al., 2004), have proven that both end-points (alterations in viability and cellular differentiation) display similar predictivity.

3.2. The whole embryo culture test (WEC)

This methodology requires an *ex vivo* exposure lasting 48 h whole rat embryos of 9.5–10 day old (Fig. 4). At this developmental stage, embryos are at a critical step of organogenesis as this is when processes such as neural tube closure, appearance of limb buds and branchial bars, heart, eye and ear development, occur. WEC is proposed in our integrated strategy as the last step because this is the test with the highest technical complexity and involves the use if animal experimentation (Fig. 2)

After exposure, embryos are scored for morphological alterations according to the parameters displayed in Table 3, which include: (a) morphology (final minus initial somite number); (b) malformations (as indicated in Table 3 from A through to R); (c) growth (crown-rump length, yolk sac diameter and head length); (d) function (yolk sac circulation, heartbeat and allantoic circulation). The final aim of this examination is to determine two end-points: the lowest concentration with no observed effect on the total morphological score ($IC_{NOEC\ TMS}$); the lowest concentration that causes the maximum malformation rate (IC_{max}). These two end-points are also used in the three discriminant functions displayed in Fig. 3 in order to classify the assessed chemical into one of the three embryotoxicants categories (strong, weak, non-embryotoxicant).

The WEC test was also subjected to an interlaboratory validation study with positive results. EURL-ECVAM determined that, like the EST, the method was scientifically validated, but is still not ready to be used for regulatory purposes (ESAC, 2002). This method is technically more complex than the EST and does not totally suppress animal experimentation. However, it allows an analysis of the effects of chemicals on later development stages and the observation of morphologically induced alterations.

WEC has been used to study the teratogenicity of lots of potential environmental contaminants, such as caffeine, methylmercury, monobutyl phthalate and methoxyacetic acid (Robinson et al., 2010); aliphatic amides used as industrial solvents like N-methyl-2-pyrrolidone (Flick et al., 2009); triazolic fungicides such as flusilazole, hexaconazole, cyproconazole, triadimefon, myclobutanil and triticonazole (de Jong et al., 2011a; Robinson et al., 2012a); and phthalates like mono(2-ethylhexyl) phthalate and monomethyl phthalate (Robinson et al., 2012b), among others.

3.3. Proposed improvements for enhancing WEC performance

A novel morphology score, called the dysmorphology score, has been proposed to graduate the intensity of the observed morphology abnormality. This scoring systems indicates a score of 5 for a normal structure, 4 for a potentially reversible abnormality or attributable to a growth delay, 3 for minor malformation, 2 for moderate malformation, 1 for severe malformation, and 0.5 for structure alterations that a gross assessment does not evidence (Augustine-Rauch et al., 2010).

Table 3
Parameters to score after exposure of rat or mouse embryos in the WEC assay. Taken from official validated protocol available in the EURL-ECVAM.

Growth parameters		Malformations (0 for normal/1 for malformed)
Yolk sac diameter (mm)		Yolk sac vessel defect
Crown-rump length (mm)		Allantois nor fused with ectoplacental cone
Head length (mm)		Allantois large size
Functional parameters (1 for normal/0 for abnormal)		Flexion deficient
		Pericardiac sac wide, filled with fluid
Yolk sac circulation		Heart ventrally turned
Allantois circulation		Posterior neuropore open
Heartbeat		Dorsal midline irregular
Somite development		Prosencephalon open
Final somite number		Rhomboencephalon narrow
Final-initial somite number		Cranial neural folds suture line irregular
Morphological scores		Head small and bent backwards
A	Yolk sac blood vessels	Craniofacial appearance abnormal
B	Allantois	Neural tube haemorrhagic
C	Flexion	Rhombencephalon large and transparent
D	Heart	Rhombencephalon narrow
E	Caudal neural tube	Otic vesicles deformed
F	Hind brain	Optic vesicles deformed
G	Mid brain	Branchial bars deformed
H	Fore brain	Maxillary process swollen
J	Otic system	Mandibular processes unapproached
K	Optic system	Mandibular process deformed
L	Olfactory system	Somites small
M	Branchial bars	Somites irregular
N	Maxillary process	Tail kinked
P	Mandibular process	Rail short and thickened
Q	Fore Limb	Subcutaneous blisters
R	Hind limb	Haemorrhages
Total morphological score (A + B + C + . . . + R)		Other

Transcriptomic approaches have been proposed as a complementary end-point, together with classical morphological end-points, to enhance the sturdiness of the WEC test. In addition, this genomic approach might also be useful for studying teratogenicity mechanisms. As expected, different chemicals were found to cause alterations in different genomic pathways. However, one common point appeared to be an alteration in the cholesterol/lipid homeostasis found in at least the following teratogens: caffeine, methyl mercury, phthalates and triazoles (Robinson et al., 2010, 2012a,b).

A streamlined WEC protocol was developed that requires the assay of a single concentration. It follows a decision tree that takes into account morphological scores of only spinal cord, heart, and the number of somite pairs (Zhang et al., 2012). This streamlined protocol reduces operational costs and animal use to 50% and yields results that does not statistically differ from the WEC validation study sponsored by EURL-ECVAM. The main problem of this procedure is that the use of a single concentration does not allow concentration-effect relationships and it is not possible to determine dose without observable adverse effect needed for risk assessment.

3.4. Acute Fish Embryotoxicity Test (FET)

Nowadays, zebrafish (*Danio rerio*) is considered an excellent friendly experimental system model. Adult simple maintenance, positive balance between results and cost, small size, high fertility rate, rapid development, optical transparency during the embryonic stage, extensive database on basic science, well-established timetable for specific developmental milestones (Fig. 5), genome completely sequenced, and the numerous commercially available kit methods where genetic or biochemical effect-response are well identified are sufficient reasons for positioning this organism as leader among the experimental systems for working on development. In addition, and according to the EU directive 2010/63/EU on the protection of animals used for scientific purposes, the earliest life-stages of animals are not defined as protected and therefore, zebrafish embryos do not fall into the regulatory frameworks dealing with animal experimentation (Yang et al., 2009; Strähle et al., 2012; Teixidó et al., 2013).

Additional advantages of this model are: (1) the high percentage of genetic control programs conserved among vertebrates and fish during development, which becomes in robustly justified the phylogenetic parallelism between fish and mammalian development; (2) the external fertilization of oviparous species as zebrafish that facilitate the handling and optical observation during the development stages (Spitsbergen and Kent, 2003); and (3) it can be used for characterize developmental phenotypes, for example using alcian blue staining for detecting craniofacial abnormalities, or acridine orange staining to track apoptotic cells during growth of the embryo (Xu et al., 2013).

The OECD accepted in July 2013 the guideline 236 (acute fish embryotoxicity (FET) test). The international regulatory acceptance of this method highlights it for testing developmental toxicity of chemicals This event makes the zebrafish a very attractive model for industry to reduce the expensive cost of the *in vivo* experiments while approaches it to the 3Rs philosophy.

All the above described advantages have positioned zebrafish test in the first stage among tests with whole embryos within our integrated strategy for assessing embryotoxicity and teratogenicity (Fig. 2).

3.5. Basic procedures in FET

Brood fish colony in a 12–16 h of photoperiod and at the maximum concentration of 1 L of water per fish is required for the test. The basic equipment needed to carry out the test is: inverted microscope, 24-well plates, an incubator to stabilize the temperature at 26 ± 1 °C and basic equipment to allow to characterize of the dilution water and stabilize the physical–chemicals parameters. Several types of spawning forms are accepted, via spawning groups or via mass spawning. The fertilized eggs, selected before cleavage of the blastodisc commence (Fig. 5) by the 16 cell-stage will be arranged individually in a 24-well plate. Five concentrations will be tested observing 100% lethality in the highest concentrations and no observable effects at the lowest concentrations tested.

A total of 168–192 fertilized eggs are used in the main test distributed as follows: 20 eggs per plate for each concentration of test, 20 eggs for the solvent control, 20 eggs for the positive control (usually 3,4-dichloroaniline), 4 eggs per plate as an internal control in each of the above plates, 24 eggs in dilution water as negative control. The assay has a maximum duration of 96 h, making observations every 24 of the following effect variables: coagulated embryos, lack of somite formation, non-detachment of the tail-bud from the yolk sac and lack of heart beat. If any of the observations made every 24 h gives a positive result is considered that the embryo is dead.

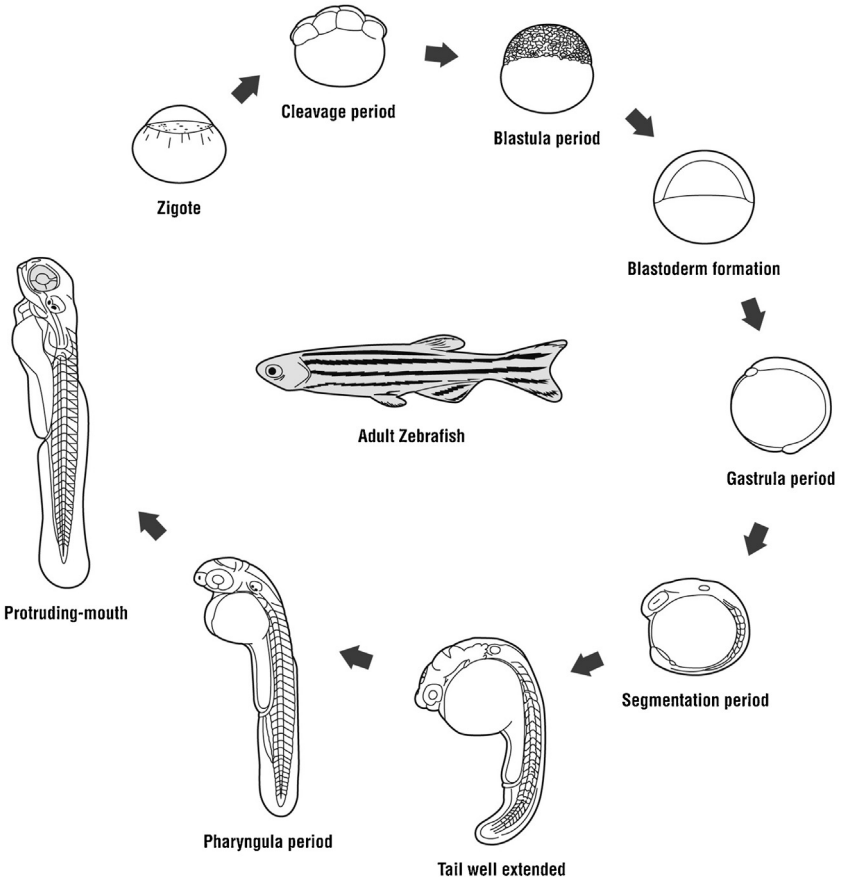


Fig. 5. Embryological stages of zebrafish. Taken form Pamies et al., 2011 (with permission).

One of the main disadvantages of this test is that cannot be applied with chemicals with molecular weight higher than 3 kDa or with bulky molecular structures. On the same line, the limited metabolic capabilities of the embryos against the juvenile or adult zebrafish individuals make recommendable to run the test also with metabolites of the parent compound (OECD Guideline 236, 2013).

However, the OECD guideline 236 is not the only one that is available as standard protocol for testing embryotoxicity with zebrafish. ISO organization accepted in 2007 standard ISO 15088:2007 (Water quality–determination of the acute toxicity of waste water to zebrafish eggs (*Danio rerio*)) in order to determine the degree of contamination of surface waters, wasted waters, municipal treated waste waters and industrial effluents.

3.6. *Zebrafish test for screen developmental toxicants*

Numerous scientific articles confirm the appropriate use of zebrafish embryos as a model to assess the safety of drugs and chemicals (Teixidó et al., 2013; McGrath and Li, 2008; Scholz et al., 2013; Pamies et al., 2011; Belanger et al., 2013; de Jong et al., 2011a). Nevertheless, it is remarkable that zebrafish test has limitations as teratogenicity test; i.e. in a screening study with 31 different chemicals (18 teratogens and 13 non-teratogens) the test was able to successfully categorize 87% of the compounds with rates of false positive and negatives of 6% (Brannen et al., 2010). However, van den Bulck et al. (2011) reported 40% rate of false positives and false negatives using a reduced set of chemicals (8 teratogens and 7 non teratogens).

Scholz et al., 2013 revealed the necessity to explore molecular targets in fish embryos because of the interesting correlations between different studies for estrogenic effects when compared to *in vivo* reproductive data.

Cheng et al. (2000) used zebrafish embryos as model system to investigate cadmium development toxicity. Effects on gastrulation, segmentation and neurogenesis processes were reported from 5 to 28 h post fertilization, obtaining a EC_{50} of 138 μM and LC_{50} at 168 μM. Molecular basis of cadmium teratogenesis effects coincided with the significant reduction in myosin heavy-chain production, alterations in the expression of the Sonic hedgehog (Shh), even-skipped-like 1 (*Eve1*) and neurotransmitter transporter-like (*Ntl*) genes that induce alterations in differentiation, apoptosis, and cell migration.

Zebrafish test has been used to evaluate the safety of ENM. Presence and exposition of ENM is an inevitable part of our daily life. Every year ENM are synthesized and released to all environmental compartments. Existence of the chorion and yolk sac, as selective barriers, seems to reduce the exposure of the embryos to ENM. Some characteristics of ENM as size, aggregate properties, dispersion or change in the surface charge density (zeta potential), combined with salts and minerals composition of the water media seems modify external components of chorion, increasing or decreasing the uptake rates and the exposure.

Transport and biocompatibility of ENM into zebrafish embryos have been studied for some nanoparticles. Silver nanoparticles are transported into embryos through chorion pore canals by Brownian diffusion from 0.19 μM concentration. Combined rates of passive diffusion and bioaccumulation of this nanomaterial showed dose-dependent abnormalities in embryo development (Lee et al., 2007).

Environmental human safety of ENM with biomedical and biotechnological applications are being studied. Pericardial edema, yolk sac edema, tail and head malformation are embryonic

malformations caused in zebrafish by exposure to silica nanoparticles. These dose-response effects of ENM that can be in contact with human tissues claim a major understanding and knowledge of biological mechanism of action (Duan et al., 2013)

However, if there is a discipline where the use of zebrafish embryos has more travel and promotion is the field of omics. Knowing the biological mode of action of environmental contaminants, such as chemicals, pesticides and other compounds with environment fate is one of the challenges of the XXI century scientific community. Approach of disciplines such as toxicogenomics, metabolomics or proteomics seem coherent alternatives to find and raise awareness relevant and useful biomarkers of exposure or effect (Link et al., 2006; Yang et al., 2007, 2009; Scholz et al., 2008; Hermsen et al., 2011). In this way, a recent study suggested that the use of individual gene expression signatures and pathway regulations may be useful for defining these biomarkers since the authors were able to, using transcriptomics, identify expressed genes mostly related to development after exposure to the embryotoxicants caffeine, carbamazepine, retinoic acid and valproic acid and the non-embryotoxic D-mannitol and saccharin (Hermsen et al., 2013).

4. Conclusions and final remarks

There are a large number of available alternatives for screening developmental toxicants that exhibit good concordance with *in vivo* results. However, a single *in vitro–ex vivo* alternative test might not be enough for identification of developmental hazards and subsequent risk assessment. These methodologies also display the same general inconveniences as any other *in vitro* test, mainly relating to possible differences in toxicokinetics and toxicodinamics, inter-species extrapolation and the *in vivo–in vitro* extrapolation of exposures. An example of this situation was reported when the potency ranking of five *p*-substituted phenols was evaluated to find that the EST identified the embryotoxic potential of phenols and provided an identical potency ranking as the WEC assay. However, the EST was unable to predict an accurate ranking for phenols as compared to their potency observed *in vivo* (Strikwold et al., 2012). The zebra fish test was the best procedure for ranking the teratogenicity of six triazole fungicides in relation to *in vivo* teratogenic potency, while WEC obtained the lowest correlation with *in vivo* results, with the EST in the middle between zebrafish and WEC (de Jong et al., 2011a).

Apart from the above-mentioned problems, it is also necessary to consider that development is a very complex biological process, and that a single *in vitro* test capable of covering this whole process with satisfactory predictivity cannot be expected. In such cases, an integrated testing strategy, as we suggest in Fig. 2, is recommended (Hartung et al., 2013a,b).

It is also remarkable that this tiered approach considers the inclusion of three different animal models (mouse, rat and zebrafish) in considerations by covering the possible inter-species variations in cellular differentiation and embryonic development, which helps minimise risks in the animal-human extrapolation. Evidently, in basis of a case-by-case study, the application of one step or more of those displayed in Fig. 2 can be avoided by following expert opinions and other available data. The tiered approach outlined in Fig. 2 has been designed for the purpose of identifying embryotoxic and teratogenic hazards, but may also prove valuable for risk assessment. Nevertheless in this case, the flow diagram cannot be followed exactly as displayed in Fig. 2 because the necessity of performing one or another test, or even the need to do an *in vivo* test, might also be determined by other considerations, such as further toxicological information available, physical properties, dose-response extrapolations and, of course, specific regulations for the material being assessed.

Acknowledgement

To Elsevier Limited for the license to reproduce Figs. 4 and 5 taken from reference Pamies et al., 2011 (see list of references).

References

Augustine-Rauch, K., Zhang, C.X., Panzica-Kelly, J.M., 2010. *In vitro* developmental toxicology assays: a review of the state of the science of rodent and zebrafish whole embryo culture and embryonic stem cell assays. Birth Defects Res. C Embryo Today 90, 87–98.

Baek, D.H., Kim, T.G., Lim, H.K., Kang, J.W., Seong, S.K., Choi, S.E., Lim, S.Y., Park, S.H., Nam, B.H., Kim, E.H., Kim, M.S., Park, K.L., 2012. Embryotoxicity assessment of developmental neurotoxicants using a neuronal endpoint in the embryonic stem cell test. J. Appl. Toxicol. 32, 617–626.

Barrier, M., Jeffay, S., Nichols, H.P., Chandler, K.J., Hoopes, M.R., Slentz-Kesler, K., Hunter 3rd., E.S., 2011. Mouse embryonic stem cell adherent cell differentiation and cytotoxicity (ACDC) assay. Reprod. Toxicol. 31, 383–391.

Belanger, S.E., Rawlings, J.M., Carr, G.J., 2013. Use of fish embryo toxicity tests for theprediction of acute fish toxicity to chemicals. Environ. Toxicol. Chem. 32, 1768–1783.

Brannen, K.C., Panzica-Kelly, J.M., Danberry, T.L., Augustine-Rauch, K.A., 2010. Development of a zebrafish embryo teratogenicity assay and quantitative prediction model. Birth Defects Res. B Dev. Reprod. Toxicol. 89, 66–77.

Buesen, R., Genschow, E., Slawik, B., Visan, A., Spielmann, H., Luch, A., Seiler, A., 2009. Embryonic stem cell test remastered: comparison between the validated EST and the new molecular FACS-EST for assessing developmental toxicity *in vitro*. Toxicol. Sci. 108, 389–400.

Buzanska, L., Sypecka, J., Nerini-Molteni, S., Compagnoni, A., Hogberg, H.T., del Torchio, R., Domanska-Janik, K., Zimmer, J., Coecke, S., 2009. A human stem cell-based model for identifying adverse effects of organic and inorganic chemicals on the developing nervous system. Stem Cells 27, 2591–25601.

Chandler, K.J., Barrier, M., Jeffay, S., Nichols, H.P., Kleinstreuer, N.C., Singh, A.V., Reif, D.M., Sipes, N.S., Judson, R.S., Dix, D.J., Kavlock, R., Hunter 3rd., E.S., Knudsen, T.B., 2011. Evaluation of 309 environmental chemicals using a mouse embryonic stem cell adherent cell differentiation and cytotoxicity assay. PLoS One 6, e18540.

Cheng, S.H., Wai, A.W.K., So, C.H., Wu, R.S.S., 2000. Cellular and molecular basis of cadmium-induced deformities in zebrafish embryos. Environ. Toxicol. Chem. 19, 3024–3031.

Colleoni, S., Galli, C., Gaspar, J.A., Meganathan, K., Jagtap, S., Hescheler, J., Sachinidis, A., Lazzari, G., 2011. Development of a neural teratogenicity test based on human embryonic stem cells: response to retinoic acid exposure. Toxicol. Sci. 124, 370–377.

de Jong, E., Barenys, M., Hermsen, S.A., Verhoef, A., Ossendorp, B.C., Bessems, J.G., Piersma, A.H., 2011a. Comparison of the mouse embryonic stem cell test, the rat whole embryo culture and the zebrafish embryotoxicity test as alternative methods for developmental toxicity testing of six 1,2,4-triazoles. Toxicol. Appl. Pharmacol. 253, 103–111.

de Jong, E., Doedée, A.M., Reis-Fernandes, M.A., Nau, H., Piersma, A.H., 2011b. Potency ranking of valproic acid analogues as to inhibition of cardiac differentiation of embryonic stem cells in comparison to their in vivo embryotoxicity. Reprod. Toxicol. 31, 375–382.

de Jong, E., Louisse, J., Verwei, M., Blaauboer, B.J., van de Sandt, J.J., Woutersen, R.A., Rietjens, I.M., Piersma, A.H., 2009. Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared with the in vivo potency of their parent compounds. Toxicol. Sci. 110, 117–124.

Di Guglielmo, C., López, D.R., De Lapuente, J., Mallafre, J.M., Suàrez, M.B., 2010. Embryotoxicity of cobalt ferrite and gold nanoparticles: a first *in vitro* approach. Reprod. Toxicol. 30, 271–276.

Duan, J., Yu, Y., Shi, H., Tian, L., Guo, C., Huang, P., Zhou, X., Peng, S., Sun, Z., 2013. Toxic effects of silica nanoparticles on zebrafish embryos and larvae. PLoS One 8, e74606.

ESAC (European Centre for Validation of Alternative Methods (ECVAM) Scientific Advisory Committee) (2002), The Use of Scientifically-Validated *In Vitro* Tests for Embryotoxicity, Available in: http://ecvam.jrc.ec.europa.eu/

Estevan, C., Vilanova, E., Sogorb, M.A., 2013. Chlorpyrifos and its metabolites alter gene expression at non-cytotoxic concentrations in D3 mouse embryonic stem cells under *in vitro* differentiation: considerations for embryotoxic risk assessment. Toxicol. Lett. 217, 14–22.

Estevan, C., Pamies, D., Sogorb, M.A., Vilanova, E., 2011. OECD guidelines and validated methods for in vivo testing of reproductive toxicity. In: Gupta, R.G. (Ed.), Reproductive and Developmental Toxicology. Academic Press, pp. 123–133.

Festag, M., Viertel, B., Steinberg, P., Sehner, C., 2007. An *in vitro* embryotoxicity assay based on the disturbance of the differentiation of murine embryonic stem cells into endothelial cells. II. Testing of compounds. Toxicol *in vitro*. 21, 1631–1640.

Flick, B., Talsness, C.E., Jäckh, R., Buesen, R., Klug, S., 2009. Embryotoxic potential of N-methyl-pyrrolidone (NMP) and three of its metabolites using the rat whole embryo culture system. Toxicol. Appl. Pharmacol. 237, 154–167.

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. Alternative Laboratory Animal 30, 151–176.

Gustafson, A.L., Stedman, D.B., Ball, J., Hillegass, J.M., Flood, A., Zhang, C.X., Panzica-Kelly, J., Cao, J., Coburn, A., Enright, B.P., Tornesi, M.B., Hetheridge, M., Augustine-Rauch, K.A., 2012. Inter-laboratory assessment of a harmonized zebrafish developmental toxicology assay–progress report on phase I. Reprod. Toxicol. 33, 155–164.

Hartung, T., Hoffmann, S., Stephens, M., 2013a. Mechanistic validation. ALTEX 30, 119–130.

Hartung, T., Luechtefeld, T., Maertens, A., Kleensang, A., 2013b. Integrated testing strategies for safety assessments. ALTEX 30, 3–18.

Hermsen, S.A., Pronk, T.E., van den Brandhof, E.J., van der Ven, L.T., Piersma, A.H., 2013. Transcriptomic analysis in the developing zebrafish embryo after compound exposure, individual gene expression and pathway regulation. Toxicol. Appl. Pharmacol. 272, 161–171.

Hermsen, S.A., Pronk, T.E., van den Brandhof, E.J., van der Ven, L.T., Piersma, A.H., 2011. Chemical class-specific gene expression changes in the zebrafish embryo after exposure to glycol ether alkoxy acids and 1,2,4-triazole antifungals. Reprod. Toxicol. 32, 245–252.

Hoelting, L., Scheinhardt, B., Bondarenko, O., Schildknecht, S., Kapitza, M., Tanavde, V., Tan, B., Lee, Q.Y., Mecking, S., Leist, M., Kadereit, S., 2013. A 3- dimensional human embryonic stem cell (hESC)-derived model to detect developmental neurotoxicity of nanoparticles. Arch. Toxicol. 87, 721–733.

Hogberg, H.T., Bressler, J., Christian, K.M., Harris, G., Makri, G., O'Driscoll, C., Pamies, D., Smirnova1, L., Wen, Z., Hartung, T., 2013. Toward a 3D model of human brain development for studying gene/environment interactions. Stem Cell Res. Ther. 4 (Suppl 1N).

Hogberg, H.T., Kinsner-Ovaskainen, A., Coecke, S., Hartung, T., Bal-Price, A.K., 2010. mRNA expression is a relevant tool to identify developmental neurotoxicants using an *in vitro* approach. Toxicol. Sci. 113, 95–115.

Hogberg, H.T., Kinsner-Ovaskainen, A., Hartung, T., Coecke, S., Bal-Price, A.K., 2009. Gene expression as a sensitive endpoint to evaluate cell differentiation and maturation of the developing central nervous system in primary cultures of rat cerebellar granule cells (CGCs) exposed to pesticides. Toxicol. Appl. Pharmacol. 235, 268–286.

Kong, D., Xing, L., Liu, R., Jiang, J., Wang, W., Shang, L., Wei, X., Hao, W., 2013. Individual and combined developmental toxicity assessment of bisphenol A and genistein using the embryonic stem cell test *in vitro*. Food Chem. Toxicol. 60, 497–505.

Lee, K.J., Nallathamby, P.D., Browning, L.M., Osgood, C.J., Xu, X.H., 2007. In vivo imaging of transport and biocompatibility of single silver nanoparticles in early development of zebrafish embryos. ACS Nano 1, 133–143.

Link, V., Shevchenko, A., Heisenberg, C.P., 2006. Proteomics of early zebrafish embryos. BMC Dev. Biol. 6, 1.

McGrath, P., Li, C.Q., 2008. Zebrafish: a predictive model for assessing drug-induced toxicity. Drug Discov. Today 13, 394–401.

Minta, M., Wilk, I., Żmudzki, J., 2004. Embryotoxicity of carbendazim in rat and hamster micromass cultures. Bull. Vet. Inst. Pulawy 48, 481–484.

Moors, M., Rockel, T.D., Abel, J., Cline, J.E., Gassmann, K., Schreiber, T., Schuwald, J., Weinmann, N., Fritsche, E., 2009. Human neurospheres as three-dimensional cellular systems for developmental neurotoxicity testing. Environ. Health Perspect. 117, 1131–1138, Erratum in: Environmental Health Perspectives. 2009. 117:A342.

Neri, T., Merico, V., Fiordaliso, F., Salio, M., Rebuzzini, P., Sacchi, L., Bellazzi, R., Redi, C.A., Zuccotti, M., Garagna, S., 2011. The differentiation of cardiomyocytes from mouse embryonic stem cells is altered by dioxin. Toxicol. Lett. 202, 226–236.

Organisation for Economic Co-operation and Development, 2013. Test No. 236: Fish Embryo Acute Toxicity (FET) Test. Available in: http://www.oecd-ilibrary.org/docserver/download/9713161e.pdf?expires=1380448247&id=id&accname=guest&checksum=C0A2EF8CA2BBE64DC412FE1FD9D100EC

Osman, A.M., van Dartel, D.A., Zwart, E., Blokland, M., Pennings, J.L., Piersma, A.H., 2010. Proteome profiling of mouse embryonic stem cells to define markers for cell differentiation and embryotoxicity. Reprod. Toxicol. 30, 322–332.

Pamies, D., Estevan, C., Sogorb, M.A., Vilanova, E., 2011. Mechanism-based models in reproductive and developmental toxicology. In: Gupta, R.G. (Ed.), Reproductive and Developmental Toxicology. Academic Press, pp. 135–216.

Panzica-Kelly, J.M., Brannen, K.C., Ma, Y., Zhang, C.X., Flint, O.P., Lehman-McKeeman, L.D., Augustine-Rauch, K.A., 2013. Establishment of a molecular embryonic stem cell developmental toxicity assay. Toxicol. Sci. 131, 447–457.

Pennings, J.L., Theunissen, P.T., Piersma, A.H., 2012. An optimized gene set for transcriptomics based neurodevelopmental toxicity prediction in the neural embryonic stem cell test. Toxicology 300, 158–167.

Piersma, A.H., Bosgra, S., van Duursen, M.B., Hermsen, S.A., Jonker, L.R., Kroese, E.D., van der Linden, S.C., Man, H., Roelofs, M.J., Schulpen, S.H., Schwarz, M., Uibel, F., van Vugt-Lussenburg, B.M., Westerhout, J., Wolterbeek, A.P., van der Burg, B., 2013. Evaluation of an alternative *in vitro* test battery for detecting reproductive toxicants. Reprod. Toxicol. 38, 53–64.

Riebeling, C., Pirow, R., Becker, K., Buesen, R., Eikel, D., Kaltenhäuser, J., Meyer, F., Nau, H., Slawik, B., Visan, A., Volland, J., Spielmann, H., Luch, A., Seiler, A., 2011. The embryonic stem cell test as tool to assess structure-dependent teratogenicity: the case of valproic acid. Toxicol. Sci. 120, 360–370.

Robinson, J.F., Tonk, E.C., Verhoef, A., Piersma, A.H., 2012a. Triazole induced concentration-related gene signatures in rat whole embryo culture. Reprod. Toxicol. 34, 275–283.

Robinson, J.F., Verhoef, A., van Beelen, V.A., Pennings, J.L., Piersma, A.H., 2012b. Dose-response analysis of phthalate effects on gene expression in rat whole embryo culture. Toxicol. Appl. Pharmacol. 264, 32–41.

Robinson, J.F., van Beelen, V.A., Verhoef, A., Renkens, M.F., Luijten, M., van Herwijnen, M.H., Westerman, A., Pennings, J.L., Piersma, A.H., 2010. Embryotoxicant-specific transcriptomic responses in rat postimplantation whole-embryo culture. Toxicol. Sci. 118, 675–685, Erratum in: Toxicological Sciences. 2011. 120:529.

Romero, A.C., Vilanova, E., Sogorb, M.A., 2011. Shortening and improving the embryonic stem cell test through the use of gene biomarkers of differentiation. J. Toxicol., 286034.

Rovida, C., Hartung, T., 2009. Re-evaluation of animal numbers and costs for in vivo tests to accomplish REACH legislation requirements for chemicals: a report by the transatlantic think tank for toxicology (t(4)). ALTEX 26, 187–208.

Russell, W.M.S., Burch, R.L., 1959. The Principles of Humane Experimental Technique. Methuen & Co. Ltd, London.

Scholz, S., Renner, P., Belanger, S.E., Busquet, F., Davi, R., Demeneix, B.A., Denny, J.S., Léonard, M., McMaster, M.E., Villeneuve, D.L., Embry, M.R., 2013. Alternatives to in vivo tests to detect endocrine disrupting chemicals (EDCs) in fish and amphibians—screening for estrogen androgen and thyroid hormone disruption. Crit. Rev. Toxicol. 43, 45–72.

Scholz, S., Fischer, S., Gündel, U., Küster, E., Luckenbach, T., Voelker, D., 2008. The zebrafish embryo model in environmental risk assessment–applications beyond acute toxicity testing. Environ. Sci. Pollut. Res. Int. 15, 394–404.

Seiler, A., Visan, A., Buesen, R., Genschow, E., Spielmann, H., 2004. Improvement of an *in vitro* stem cell assay for developmental toxicity: the use of molecular endpoints in the embryonic stem cell test. Reprod. Toxicol. 18, 231–240.

Spielmann, H., Seiler, A., Bremer, S., Hareng, L., Hartung, T., Ahr, H., Faustman, E., Haas, U., Moffat, G.J., Nau, H., Vanparys, P., Piersma, A., Sintes, J.R., Stuart, J., 2006. The practical application of three validated *in vitro* embryotoxicity tests. The report and recommendations of an ECVAM/ZEBET workshop (ECVAM workshop 57). Altern. Lab. Anim. 34, 527–538.

Spitsbergen, J.M., Kent, M.L., 2003. The state of the art of the zebrafish model for toxicology and toxicologic pathology research–advantages and current limitations. Toxicol. Pathol. 31, 62–87.

Strähle, U., Scholz, S., Geisler, R., Greiner, P., Hollert, H., Rastegar, S., Schumacher, A., Selderslaghs, I., Weiss, C., Witters, H., Braunbeck, T., 2012. Zebrafish embryos as an alternative to animal experiments–a commentary on the definition of the onset of protected life stages in animal welfare regulations. Reprod. Toxicol. 33, 128–132.

Strikwold, M., Woutersen, R.A., Spenkelink, B., Punt, A., Rietjens, I.M., 2012. Relative embryotoxic potency of *p*-substituted phenols in the embryonic stem cell test (EST) and comparison to their toxic potency in vivo and in the whole embryo culture (WEC) assay. Toxicol. Lett. 213, 235–242.

Suzuki, N., Ando, S., Yamashita, N., Horie, N., Saito, K., 2011. Evaluation of novel high-throughput embryonic stem cell tests with new molecular markers for screening embryotoxic chemicals *in vitro*. Toxicol. Sci. 124, 460–471.

Suzuki, N., Yamashita, N., Koseki, N., Yamada, T., Kimura, Y., Aiba, S., Toyoizumi, T., Watanabe, M., Ohta, R., Tanaka, N., Saito, K., 2012. Assessment of technical protocols for novel embryonic stem cell tests with molecular markers (Hand1- and Cmya1-ESTs): a preliminary cross-laboratory performance analysis. J. Toxicol. Sci. 37, 845–851.

Teixidó, E., Piqué, E., Gómez-Catalán, J., Llobet, J.M., 2013. Assessment of developmental delay in the zebrafish embryo teratogenicity assay. Toxicol *In Vitro.* 27, 469–478.

Theunissen, P.T., Pennings, J.L., van Dartel, D.A., Robinson, J.F., Kleinjans, J.C., Piersma, A.H., 2013. Complementary detection of embryotoxic properties of substances in the neural and cardiac embryonic stem cell tests. Toxicol. Sci. 132, 118–130.

Theunissen, P.T., Robinson, J.F., Pennings, J.L., de Jong, E., Claessen, S.M., Kleinjans, J.C., Piersma, A.H., 2012a. Transcriptomic concentration-response evaluation of valproic acid, cyproconazole, and hexaconazole in the neural embryonic stem cell test (ESTn). Toxicol. Sci. 125, 430–438.

Theunissen, P.T., Robinson, J.F., Pennings, J.L., van Herwijnen, M.H., Kleinjans, J.C., Piersma, A.H., 2012b. Compound-specific effects of diverse neurodevelopmental toxicants on global gene expression in the neural embryonic stem cell test (ESTn). Toxicol. Appl. Pharmacol. 262, 240–330.

Theunissen, P.T., Schulpen, S.H., van Dartel, D.A., Hermsen, S.A., van Schooten, F.J., Piersma, A.H., 2010. An abbreviated protocol for multilineage neural differentiation of murine embryonic stem cells and its perturbation by methyl mercury. Reprod. Toxicol. 29, 383–392.

van Dartel, D.A., Pennings, J.L., de la Fonteyne, L.J., Brauers, K.J., Claessen, S., van Delft, J.H., Kleinjans, J.C., Piersma, A.H., 2011a. Evaluation of developmental toxicant identification using gene expression profiling in embryonic stem cell differentiation cultures. Toxicol. Sci. 119, 126–134.

van Dartel, D.A., Pennings, J.L., Robinson, J.F., Kleinjans, J.C., Piersma, A.H., 2011b. Discriminating classes of developmental toxicants using gene expression profiling in the embryonic stem cell test. Toxicol. Lett. 201, 143–151.

van Dartel, D.A., Pennings, J.L., de la Fonteyne, L.J., van Herwijnen, M.H., van Delft, J.H., van Schooten, F.J., Piersma, A.H., 2010a. Monitoring developmental toxicity in the embryonic stem cell test using differential gene expression of differentiation-related genes. Toxicol. Sci. 116, 130–139.

van Dartel, D.A., Pennings, J.L., van Schooten, F.J., Piersma, A.H., 2010b. Transcriptomics-based identification of developmental toxicants through their

interference with cardiomyocyte differentiation of embryonic stem cells. Toxicol. Appl. Pharmacol. 243, 420–428.
van Dartel, D.A., Pennings, J.L., Hendriksen, P.J., van Schooten, F.J., Piersma, A.H., 2009. Early gene expression changes during embryonic stem cell differentiation into-cardiomyocytes and their modulation by monobutyl phthalate. Reprod. Toxicol. 27, 93–102.
van den Bulck, K., Hill, A., Mesens, N., Diekman, H., De Schaepdrijver, L., Lammens, L., 2011. Zebrafish developmental toxicity assay: a fishy solution to reproductive toxicity screening, or just a red herring? Reprod. Toxicol. 32, 213–219.
Wang, Y., Fan, Y., Puga, A., 2010. Dioxin exposure disrupts the differentiation of mouse embryonic stem cells into cardiomyocytes. Toxicol. Sci. 115, 225–237.
Xu, B., Lee, K.K., Zhang, L., Gerton, J.L., 2013. Stimulation of mTORC1 with L-leucine rescues defects associated with Roberts syndrome. PLoS Genet. 9, e1003857.
Yang, L., Ho, N.Y., Alshut, R., Legradi, J., Weiss, C., Reischl, M., Mikut, R., Liebel, U., Müller, F., Strähle, U., 2009. Zebrafish embryos as models for embryotoxic and teratological effects of chemicals. Reprod. Toxicol. 28, 245–253.
Yang, L., Kemadjou, J.R., Zinsmeister, C., Bauer, M., Legradi, J., Müller, F., Pankratz, M., Jäkel, J., Strähle, U., 2007. Transcriptional profiling reveals barcode-like toxicogenomic responses in the zebrafish embryo. Genome Biol. 8, R227.
Zhang, C., Cao, J., Kenyon, J.R., Panzica-Kelly, J.M., Gong, L., Augustine-Rauch, K., 2012. Development of a streamlined rat whole embryo culture assay for classifying teratogenic potential of pharmaceutical compounds. Toxicol. Sci. 127, 535–546.