Fundamental Toxicological Sciences (Fundam. Toxicol. Sci.)
Vol.3, No.2, 79-87, 2016

Original Article

Comparative study of the zebrafish embryonic toxicity test and mouse embryonic stem cell test to screen developmental toxicity of human pharmaceutical drugs

Atsuto Inoue[1,*], Yuhei Nishimura[2,3,4,5,6,*], Norihito Matsumoto[1], Noriko Umemoto[2,3,4,5,6], Yasuhito Shimada[2,3,4,5,6], Toru Maruyama[1], Kana Kayasuga[1], Motohiko Morihara[1], Jun Katagi[1], Tsutomu Shiroya[1], Yasushi Hirota[1], Soonih Kim[1] and Toshio Tanaka[2,3,4,5,6]

[1]Ono Pharmaceutical Co, Ltd., 3-1-1, Sakurai, Shimamoto-cho, Mishimagun, Osaka 618-8585, Japan
[2]Department of Molecular and Cellular Pharmacology, Pharmacogenomics and Pharmacoinformatics, Mie University Graduate School of Medicine, 2-174, Edobashi, Tsu, Mie 514-8507, Japan
[3]Mie University Medical Zebrafish Research Center, 2-174, Edobashi, Tsu, Mie 514-8507, Japan
[4]Department of Systems Pharmacology, Mie University Graduate School of Medicine, 2-174, Edobashi, Tsu, Mie 514-8507, Japan
[5]Department of Omics Medicine, Mie University Industrial Technology Innovation Institute, 2-174, Edobashi, Tsu, Mie 514-8507, Japan
[6]Department of Bioinformatics, Mie University Life Science Research Center, 2-174, Edobashi, Tsu, Mie 514-8507, Japan

(Received March 11, 2016; Accepted March 18, 2016)

ABSTRACT — According to International Conference on Harmonization guidelines, each drug in development for administration to women of child-bearing potential must be tested for possible developmental toxicities using at least two species (a rodent and non-rodent). With the high cost and slow pace of embryonic-fetal toxicity testing in mammals, both the zebrafish embryonic toxicity test (ZET) and mouse embryonic stem cell test (mEST) have been shown to be useful to assess developmental toxicity of various chemical compounds, including human pharmaceutical drugs, in a high-throughput manner. However, comparative study of the sensitivity and specificity of these methods using the same set of human pharmaceutical drugs is scarce. In this study, we assessed developmental toxicity tests of 39 chemical compounds, including human pharmaceutical drugs, in both the ZET and mEST. The accuracy, sensitivity, and specificity of the ZET were 69%, 59%, and 82%, respectively. The accuracy, sensitivity, and specificity of the mEST were 64%, 50%, and 82%, respectively. As a result, both the ZET and mEST showed acceptable accuracies compared with rat embryo-fetal toxicity study and Food and Drug Administration pregnancy categories. By comparing the results between the ZET and mEST, we identified different types of true positives and true negatives. Thus, complementary tests using both the ZET and mEST may better predict the developmental toxicity of human pharmaceuticals.

Key words: Zebrafish, Embryonic stem cell test, Developmental toxicity

INTRODUCTION

According to International Conference on Harmonization (ICH) guidelines (ICH, 2005, 2015), each pharmaceutical drug in development for administration to women of child-bearing potential must be tested for possible developmental toxicities using at least two species (a rodent and non-rodent). Traditional toxicity testing requires collecting data on one chemical at a time using animal models such as the rat and rabbit. Achieving registration of a drug requires the use of a high number of animals for teratogenicity safety assessment, which is not in line with the current trend of applying the replacement, reduction, and refinement of animal studies (3R) principle

Correspondence: Toshio Tanaka (E-mail: tanaka@doc.medic.mie-u.ac.jp)
*These authors equally contributed to this work.

(Nishimura *et al*., 2016).

For example, a whole embryonic culture system is one of the candidates for the alternative methods, which shows high predictivity (Genschow *et al*., 2002). However, it requires many animals and is thus not favored from the 3R viewpoint. Additionally, numerous chemical compounds would be applied. Therefore, it is not recommended for many drug candidates in the early drug-screening stage. To date, the zebrafish embryonic toxicity test (ZET) and mouse embryonic stem cell test (mEST) have been shown to be useful to assess developmental toxicity of various chemicals including human pharmaceutical drugs (Nishimura *et al*., 2016; Matsumoto, 2009). In general, smaller amounts of compounds (less than 100 mg for each test) are used in both the ZET and mEST compared with rat embryo-fetal toxicity studies. However, comparative study of the accuracy, sensitivity, and specificity of these methods using the same set of human pharmaceutical drugs is scarce. In the present study, we assessed developmental toxicity tests of 39 chemicals in both the ZET and mEST, and show that it is possible to detect the developmental toxicities at the drug-screening stage using both tests.

MATERIALS AND METHODS

Compounds and cells

We assessed the ZET and mEST using 22 teratogens (atorvastatin, busulfan, clarithromycin, cytarabine, dasatinib, dexamethasone, diclofenac, diphenhydramine, fluticasone, hydroxyurea, imatinib, indomethacin, methotrexate, retinoic acid, ribavirin, thalidomide, theophylline, thio-TEPA, topiramate, valproic acid and 2 in-house compounds; ONO-A and ONO-B) and 17 non-teratogens (amlodipine, ascorbate, cetirizine, dipyridamole, famotidine, fexofenadine, lansoprazole, montelukast, omeprazole, penicillin G, saccharin, sodium cyclamate, ticlopidine, valpromide, zafirlukast and 2 in-house ONO compounds; ONO-C and ONO-D).

ONO-A, ONO-B, ONO-C, and ONO-D were synthesized by ONO Pharmaceutical (Osaka, Japan). Ascorbate, busulfan, cetirizine, cytarabine, diphenhydramine, famotidine, fexofenadine, fluticasone, hydroxyurea, lansoprazole, omeprazole, penicillin G, ribavirin, theophylline, thio-TEPA, ticlopidine, topiramate, and valproic acid were purchased from Sigma-Aldrich Japan (Tokyo, Japan). Diclofenac, dipyridamole, indomethacin, and methotrexate were purchased from Microsource (Gaylordsville, CT, USA). Dasatinib, imatinib, montelukast, and zafirlukast were purchased from Sequoia Research Product (Pangbourne, UK). Amlodipine and dimethyl sulfoxide (DMSO) were purchased from Kanto Chemicals (Tokyo, Japan). Atorvastatin was purchased from Chempacific (Baltimore, MD, USA). Clarithromycin and valpromide were purchased from Wako Pure Chemicals (Osaka, Japan). Dexamethasone was purchased from Enzo (New York, NY, USA). Retinoic acid and saccharin were purchased from Tokyo Chemical Industry (Tokyo, Japan). Sodium cyclamate was purchased from Supelco (Bellefonte, PA, USA). Thalidomide was purchased from Tocris (Minneapolis, MN, USA). D3 mouse embryonic stem cells (ES-D3, CRL-1934) and 3T3 mouse embryonic fibroblasts (3T3 fibroblasts, CL A31) were purchased from the American Type Culture Collection (Manassas, VA, USA) and DS Pharma Biomedical (Osaka, Japan), respectively.

ZET

Zebrafish (AB-line, 4 months old) were mated, and the fertilized eggs were used for the assay without dechorionation. Eggs were incubated with the compounds in fish medium (2 mM $CaCl_2$, 0.5 mM $MgSO_4$, 0.7 mM $NaHCO_3$, and 0.07 mM KCl, pH 7.4) for 80 hours post-fertilization (hpf). The fish medium was changed every day.

Chemical compounds were dissolved in 100% DMSO and adjusted to 10 mM concentrations. The final concentration of DMSO of the highest drug concentration in the ZET was 1%. Morphological evaluations of the embryos were performed at 80 hpf using a Nikon SMZ800 stereoscopic microscope. Teratogenic effects were recorded in terms of the following endpoints: malformation of the head, malformation of sacculi/otoliths, malformation of the tail, malformation of the heart, pericardial edema, yolk deformation, and scoliosis. Chemicals that showed an anomaly in at least one teratogenic endpoint at 80 hpf were judged as teratogens in the ZET. Compounds that showed anomalies without death were judged as ZET positive. The compounds without anomalies at 100 μM were judged as ZET negative.

mEST

The mEST was performed according to the European Centre for the Validation of Alternative Methods (ECVAM) international validation study (Genschow *et al*., 2002; Matsumoto, 2009; Riebeling *et al*., 2015). Briefly, ES-D3 cells and 3T3 fibroblasts were treated with chemicals for 10 days. Subsequently, ES-D3 cells that were differentiated toward heart development via embryoid bodies were scored for beating embryoid body outgrowths. Cytotoxic effects on ES-D3 cells and 3T3 fibroblasts were determined using an MTT (3-(4,5-di-methylthiazol-2-yl)- 2,5-diphenyl tetrazolium bromide) assay

(Mosmann, 1983). Chemical compounds were dissolved in 100% DMSO. The final concentration of DMSO in the mEST was adjusted to 1%.

The experimental parameters, prediction models, and classification criteria of the mEST are shown in Table 1. Class 3 chemicals, and Class 2 chemicals with a score calculated by subtracting linear discriminant functions I from II is higher than 4 were judged as mEST positive (Matsumoto, 2009). Class 2 chemicals with a score calculated by subtracting linear discriminant functions I from II is equal to or lower than 4 and Class 1 chemicals were judged as mEST negative (Matsumoto, 2009).

RESULTS

Typical images of zebrafish exposed to the vehicle, retinoic acid, montelukast, zafirlukast, omeprazole, lansoprazole, theophylline, dipyridamole, thio-TEPA, busulfan, dasatinib, or imatinib are shown in Fig. 1. The results of the ZET are shown in Table 2. Among the 39 compounds evaluated in this study, the accuracy, sensitivity, and specificity of the ZET were 69%, 59%, and 82%, respectively (Fig. 2). The accuracy, sensitivity, and specificity of the mEST were 64%, 50%, and 82%, respectively (Fig. 3).

Positivity in the ZET

True positive compounds in the ZET were as follows: atorvastatin, dasatinib, dexamethasone, diclofenac, indomethacin, methotrexate, ONO-A, ONO-B, theophylline, thio-TEPA, topiramate, retinoic acid, and valproic acid.

False positive compounds in the ZET were amlodipine, montelukast, and ticlopidine.

Negativity in the ZET

True negative compounds in the ZET were as follows: ascorbate, cetirizine, dipyridamole, famotidine, fexofenadine, lansoprazole, ONO-C, ONO-D, omeprazole, penicillin G, saccharin, sodium cyclamate, valpromide, and zafirlukast.

False negative compounds in the ZET were as follows: busulfan, clarithromycin, cytarabine, diphenhydramine, fluticasone, hydroxyurea, imatinib, ribavirin, and thalidomide.

Positivity in the mEST

True positive compounds in the mEST were as follows: atorvastatin, busulfan, cytarabine, dexamethasone, diclofenac, diphenhydramine, hydroxyurea, imatinib, methotrexate, retinoic acid, and thio-TEPA.

False positive compounds in mEST were dipyridamole and ticlopidine.

Negativity in the mEST

True negative compounds in the mEST were as follows: amlodipine, ascorbate, cetirizine, famotidine, fexofenadine, lansoprazole, montelukast, omeprazole, ONO-C, ONO-D, penicillin G, saccharin, valpromide, and zafirlukast.

False negative compounds in the mEST were as follows: clarithromycin, dasatinib, fluticasone, indomethacin, ONO-A, ONO-B, ribavirin, theophylline, topiramate, and valproic acid.

Not judged in the mEST

In the mEST, we could not judge the teratogenicity of sodium cyclamate or thalidomide. The parameters of the mEST (Different_ES, Tox_ES, and Tox_3T3) for these

Table 1. Experimental parameters, prediction models, and classification criteria of the mEST (Genschow *et al.,* 2002).

Experimental parameters of mEST assays:
1) 50% inhibition concentration of the differentiation of ES-D3 cell (Different_ES)
2) 50% inhibition concentration of the proliferation of ES-D3 cell (Tox_ES)
3) 50% inhibition concentration of the proliferation of 3T3 fibroblast (Tox_3T3)
The prediction models of mEST:
I : 5.916*Log (Tox_3T3) + 3.500*Log (Tox_ES) - 5.307*[(Tox_3T3 - Different_ES) / Tox_3T3] - 15.27
II : 3.651*Log (Tox_3T3) + 2.394*Log (Tox_ES) - 2.033*[(Tox_3T3 - Different_ES) / Tox_3T3] - 6.85
III : -0.125*Log (Tox_3T3) - 1.917*Log (Tox_ES) + 1.500*[(Tox_3T3 - Different_ES) / Tox_3T3] - 2.67
Classification criteria:
Class 1 Non-embryotoxic, If I>II and I>III
Class 2 Weak embryotoxic, If II>I and II>III
Class 3 Strongly embryotoxic, If III>I and III>II

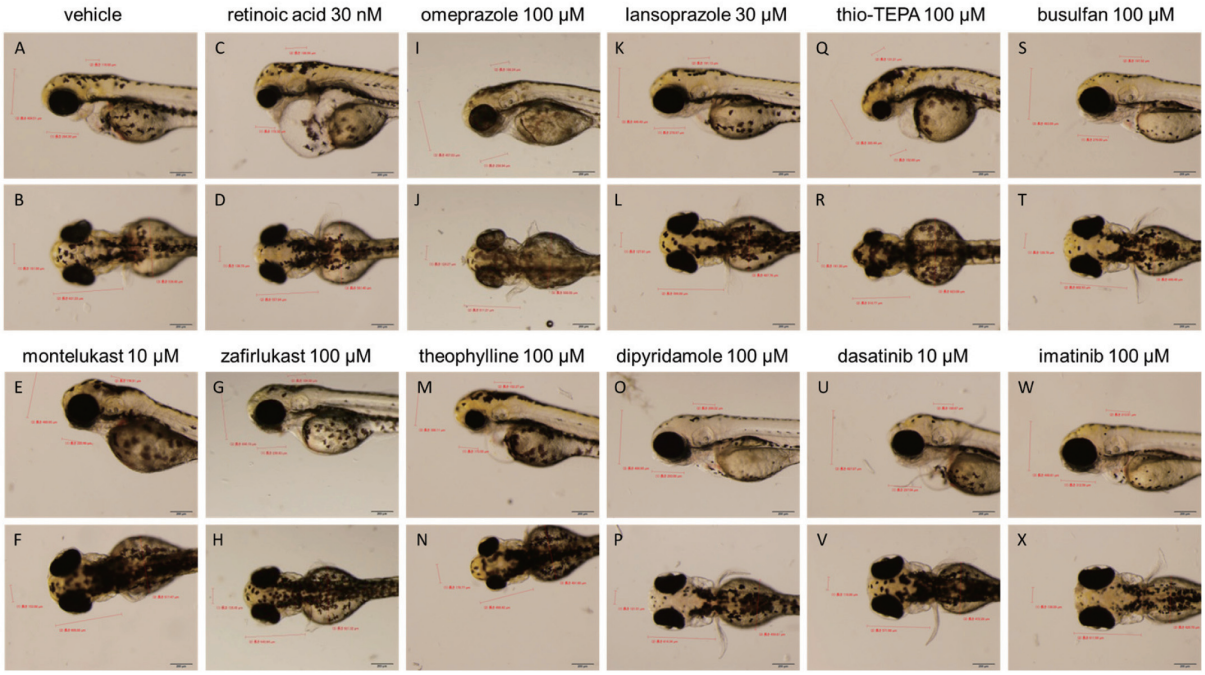

Fig. 1. Typical images of zebrafish exposed to various compounds. A) B) vehicle. C) D) retinoic acid. E) F) montelukast. G) H) zafirlukast. I) J) omeprazole. K) L) lansoprazole. M) N) theophylline. O) P) dipyridamole. Q) R) thio-TEPA. S) T) busulfan. U) V) dasatinib. W) X) imatinib.

compounds could not be calculated because the effects of these compounds on mES cells and 3T3 fibroblasts were too weak.

DISCUSSION

In this study, the accuracy, sensitivity, and specificity of the ZET were 69%, 59%, and 82% (Fig. 2), and those of the mEST were 64%, 50%, and 82% (Fig. 3), respectively. These results suggest that the ability to detect the developmental toxicities of chemical compounds is almost the same between the ZET and mEST under our assay conditions.

In other reports, the accuracies of the ZET (87% and 90%) (Brannen *et al*., 2010; Yamashita *et al*., 2014) and mEST (78% and 100%) (Genschow *et al*., 2002; Riebeling *et al*., 2015) were higher than those in our study. The reason for the lower accuracy in our assay system is unclear. One of the possible reasons for the low sensitivity in the ZET (59%) is the hydrophobicity of the chemical compounds tested in the present study. The teratogenicity of hydrophilic drugs, including hydroxyurea (logP:-1.8), busulfan (logP:-0.9), and cytarabine (logP: -2.8), could not be detected in the ZET. These compounds are cytotoxic anti-cancer drugs that have been judged as positive in the mEST and showed anomalies in rat embryonic-fetal toxicity studies, suggesting the limitation of the ZET. It would be better to predict the developmental toxicities more precisely using both the ZET and mEST, especially for hydrophilic compounds. Additionally, montelukast showed remarkable changes and was judged as positive in the ZET. The logP value of montelukast is higher (logP:7.3) than that of other compounds. We hypothesize that a higher exposure concentration of montelukast to zebrafish might affect the morphology of zebrafish. Further studies are needed to elucidate the relationship between logP and ZET positive/negative judgements.

In the ZET, one reason for the low sensitivity (59%) and high specificity (82%) is likely the penetration of chorions. Nishimura *et al*. (2016) indicated that the penetration of chemicals through the chorion can vary depending on the physicochemical properties, cationic charge concentration in the medium, and electrostatic attraction between chemicals and the chorion. For example, we could not detect the embryonic toxicity of thalidomide in the ZET or mEST. Another study could detect the embryonic toxicity of 400 μM thalidomide in the ZET with

Table 2. Comparison between the results from the rat developmental toxicity test, ZET, and mEST of 39 chemical compounds including human pharmaceutical drugs.

Compound	CAS No.	Mode of Action	MW	logP	FDA Categories	Rat embryo -fetal Tox	ZET	ZET LC_{50} (μM)	mEST (**Class**, II-I)	Diff ES (IC_{50}:μM)	Tox ES (IC_{50}:μM)	Tox 3T3 (IC_{50}:μM)
cetirizine	83881-51-0	antagonist of histamine H1 receptor	389	2.8	B	N	N	>100	N (**Class2**, 3.98)	57	368	321
diphenhydramine	58-73-1	antagonist of histamine H1 receptor	255	3.4	B	-	N	>100	P (**Class2**, 6.02)	33	66	157
famotidine	76824-35-6	antagonist of histamine H2 receptor	337	-0.6	B	N	N	>100	N (**Class2**, 2.18)	148	1426	1746
montelukast	158966-92-8	antagonist of leukotriene D4 receptor	586	7.3	B	N	P	30	N (**Class2**, 3.17)	44	70	75
zafirlukast	107753-78-6	antagonist of leukotriene D4 receptor	576	4.8	B	N	N	>100	N (**Class2**, 0.39)	163	81	97
ticlopidine	55142-85-3	blocker of ADP receptor	264	4.3	B	N	P	100	P (**Class2**, 6.79)	28	17	76
lansoprazole	103577-45-3	inhibitior of the H+/K+ ATPase	369	2.8	B	N	N	>100	N (**Class2**, 1.40)	101	97	60
dipyridamole	58-32-2	inhibitor of phosphodiesterase	505	1.5	B	N	N	>100	P (**Class2**, 6.08)	12	16	16
penicillin G	61-33-6	antibiotics	334	-	B	-	N	>100	N (**Class1**)	4800	2368	2476
dexamethasone	50-02-2	agonist of glucocorticoid receptor	392	1.8	C	P	P	>100	P (**Class2**, 4.12)	96	45	138
fluticasone	397864-44-7	agonist of glucocorticoid receptor	539	3.7	C	P	N	>100	N (**Class2**, 1.28)	22	24	9.1
topiramate	97240-79-4	antagonist of AMPA/kainate receptor	339	1.3	C	P	P	>100	N (**Class2**, 3.58)	393	936	2357
fexofenadine	83799-24-0	antagonist of histamine H1 receptor	502	5.0	C	N	N	>100	N (**Class2**, 2.49)	42	744	1135
amlodipine	88150-42-9	blocker of voltage-gated L-type Ca^{2+} channel	409	2.2	C	N	P	>100	N (**Class2**, 0.31)	6.2	7.9	1.9
omeprazole	73590-58-6	inhibitior of the H^+/K^+ ATPase	345	1.7	C	N	N	>100	N (**Class2**, 3.81)	90	119	122
clarithromycin	81103-11-9	inhibitor of bacterial protein synthesis	748	3.2	C	P	N	>100	N (**Class2**, 3.29)	17	189	551
indomethacin	53-86-1	inhibitor of cyclooxygenase	358	4.3	C	P	P	>100	N (**Class2**, 2.86)	156	154	183
theophylline	58-55-9	inhibitor of phosphodiesterase, HDAC activator	180	-0.3	C	P	P	>100	N (**Class2**, 2.80)	500	471	978
ascorbate	50-81-7	scavengers of free radicals	176	-1.6	C	N	N	>100	N (**Class1**)	1932	277	56
thio-TEPA	52-24-4	alkylating agent	189	0.2	D	P	P	>100	P (**Class3**)	0.64	0.12	1.6
busulfan	55-98-1	alkylating agent	246	-0.9	D	P	N	>100	P (**Class3**)	20	6.1	30
cytarabine	147-94-4	inhibitor of DNA synthesis	243	-2.8	D	P	N	>100	P (**Class3**)	0.27	0.038	0.16
diclofenac	15307-86-5	inhibitor of cyclooxygenase	296	4.5	D	N	P	50	P (**Class2**, 4.08)	85	268	163
valproic acid	99-66-1	inhibitor of GABA transaminase	144	2.5	D	P	P	>100	N (**Class2**, 3.22)	327	802	2415
hydroxyurea	127-07-1	inhibitor of ribonucleoside diphosphate reductase	76	-1.8	D	P	N	>100	P (**Class3**)	40	17	51
dasatinib	302962-49-8	inhibitor of tyrosine kinase	488	1.8	D	P	P	>100	N (**Class1**)	9.2	1.6	1.6
imatinib	152459-95-5	inhibitor of tyrosine kinase	494	3.5	D	P	N	>100	P (**Class2**, 6.02)	17	11	23
atorvastatin	134523-00-5	inhibitor of HMG-CoA reductase	558	4.4	X	N	P	20	P (**Class3**)	1.6	6.2	5.2
ribavirin	36791-04-5	inhibitor of nucleic acid synthesis	244	-1.9	X	P	N	>100	N (**Class2**, 3.42)	20	6.2	12
methotrexate	59-05-2	inhibitor of tetrahydrofolate dehydrogenase	454	-0.9	X	P	P	>100	P (**Class3**)	0.042	0.035	0.031
retinoic acid	302-79-4	agonist of RAR and RXR	300	6.7	X	P	P	0.2	P (**Class3**)	0.00087	0.0056	12
thalidomide	50-35-1	anticancer	258	0.5	X	N	N	>100	N (ND)	>2711	>2711	>2711
ONO-A	-	GPCR antagonist	495	4.5	-	P	P	>100	N (**Class2**, 2.18)	88	111	71
ONO-B	-	GPCR antagonist	501	5.8	-	P	P	>100	N (**Class2**, 2.71)	106	190	128
ONO-C	-	GPCR antagonist	471	5.8	-	N	N	>100	N (**Class2**, 2.41)	148	419	225
ONO-D	-	GPCR antagonist	440	6.7	-	N	N	30	N (**Class1**)	270	412	141
sodium cyclamate	139-05-9	an artificial sweetener	201	-	-	N	N	>100	N (ND)	>2485	>2485	>2485
saccharin	81-07-2	an artificial sweetener	183	0.7	-	N	N	>100	N (**Class1**)	16572	5524	19990
valpromide	2430-27-5	inhibitor of GABA transaminase	143	1.8	-	N	N	>100	N (**Class2**, 0.62)	2063	3097	3167

N: negative; P: positive; ND: not determined.

Teratology risk	ZET negative	ZET positive
Low	True Negative: 14 compounds (82%) ONO-C, ONO-D, Valpromide, Zafirlukast, Cetirizine, Dipyridamole, Omeprazole, Saccharin, Penicillin G, Ascorbate, Famotidine, Fexofenadine, Lansoprazole, Sodium cyclamate	False Positive: 3 compounds (18%) Montelukast, Ticlopidine, Amlodipine
High	False Negative: 9 compounds (41%) Clarithromycin, Hydroxyurea, Busulfan, Cytarabine, Thalidomide, Diphenhydramine, Fluticasone, Ribavirin, Imatinib	True Positive: 13 compounds (59%) ONO-A, ONO-B, Indometacin, Atorvastatin, Methotrexate, Dexamethasone, Valproic acid, Diclofenac, Theophylline, Thio-TEPA, Topiramate, Dasatinib

Fig. 2. Accuracy, sensitivity and specificity of the ZET.

mEST results

Teratology risk	Class 1 (II−I <4.0)	Class 2 (II−I >4.0)	Class 3
Low	True Negative: 14 compounds (82%) ONO-D, Saccharin, Penicillin G, Ascorbate Zafirlukast, Valpromide, ONO-C, Montelukast, Omeprazole, Amlodipine, Lansoprazole, Famotidine, Fexofenadine, Cetirizine	False positive: 2 compounds (12%) Dipyridamole, Ticlopidine	
High	False Negative: 10 compounds (45%) Dasatinib ONO-A, ONO-B, Theophylline, Indometacin, Valproic acid, Fluticasone, Clarithromycin, Ribavirin, Topiramate	True Positive: 11 compounds (50%) Dicrofenac, Dexamethasone, Imatinib, Diphenhydramin Atorvastatin, Methotrexate, Retinoic acid, Busulfan, Cytarabine, Hydroxyurea, Thio-TEPA	

Not judged: Sodium cyclamate, Thalidomide

Fig. 3. Accuracy, sensitivity, and specificity of the mEST.

dechorionation (Ito *et al*., 2010). One of the reasons why we could not detect the toxicity of thalidomide in the ZET might be dechorionation. Further studies would be needed to explain these discrepancies.

Another possible reason for the low sensitivity in the ZET (59%) is the highest concentration in our assay system. Organisation for Economic Co-operation and Development (OECD) TG236 recommends the highest concentration that causes 100% lethality and the lowest concentration that causes no observable effect (OECD, 2013). For drug screening at the early exploratory stages, we could not prepare large amounts of chemical compounds. Thus, we evaluated the compounds at 100 μM in the ZET as the highest concentration. The highest concentration in a ZET by Yamashita *et al*. (2014) was 1,000 μM. Although the set of chemical compounds is not same in our assay system, they showed higher sensitivity (88%). Furthermore, both we and Yamashita *et al*. (2014) used the same seven compounds (dexamethasone, diclofenac, indomethacin, methotrexate, penicillin G, retinoic acid, and ribavirin), and the results were the same without indomethacin. Additionally, we evaluated hydroxyurea at 10 mM and thalidomide at 1 mM, but they did not cause any morphological changes in the ZET (data not shown). At present, it is unclear whether the highest concentration affected the low sensitivity of ZET results.

In the mEST, one of the reasons for the low sensitivity (50%) and high specificity (82%) is likely the limitation of the method itself. The differentiation value is only obtained from cardiomyocyte differentiation, thereby lacking detection of the effects of compounds on organs other than the heart (De Jong *et al*., 2014). This single differentiation endpoint might limit the predictive value of our assay. Another possible reason is using mouse-derived cells. For example, we could not detect the teratogenicity of thalidomide in the mEST. Recently, Aikawa *et al*. (2014) demonstrated the embryonic toxicity of thalidomide (weak teratogenicity) using human induced pluripotent stem cells. Further modification of the protocol might be needed to achieve better predictions in the mEST.

We also tested in-house compounds ONO-A, B, C, and D (Table 2) to evaluate the generality and capability of the ZET and mEST as teratogenicity assessment systems. ONO-A and ONO-B were precisely determined as teratogenic substances in the ZET (Fig. 2) in accordance with the results of a rat embryonic-fetal toxicity study (data not shown). In contrast, prediction of the teratogenicity of ONO-A and ONO-B failed in the mEST (Fig. 3). These results indicate that the ZET is more suitable for assessment of these compounds. We do not understand why these two tests provided different results at this time. Further studies are needed to clarify the limitation of these assessments.

A Venn diagram of the results from the ZET, mEST, and rat embryonic-fetal toxicity study using 31 compounds including human pharmaceutical drugs is shown in Fig. 4. Among the Food and Drug Administration (FDA) pregnancy category B compounds (Fig. 4, upper box), seven of nine were judged as non-teratogens in the rat embryonic-fetal toxicity study (two of nine are unknown), and seven of nine in the ZET and six of nine in the mEST were judged as non-teratogens. Among FDA pregnancy category C compounds (Fig. 4, middle box), four of 10 in the rat embryonic-fetal toxicity study, five of 10 in the ZET, and nine of 10 in the mEST were judged as non-teratogens. Among FDA pregnancy category D and X compounds (Fig. 4, lower box), 10 of 13 in the rat embryonic-fetal toxicity study, seven of 13 in the ZET, and nine of 13 in the mEST were judged as teratogens. In the present study, both the ZET and mEST showed acceptable accuracies and correlations with the rat embryonic-fetal toxicity study and FDA pregnancy categories.

Among FDA pregnancy category B compounds, two in the ZET, and three in the mEST were judged as positive. In particular, ticlopidine was positive in both the ZET and mEST, but negative in the rat embryonic-fetal toxicity study. It is known that ticlopidine is metabolized into an active form *in vivo* (Hagihara *et al*., 2008). Therefore, the differences in metabolism between the rat (*in vivo*), zebrafish (*in vivo*), and mouse (*in vitro*) might affect the results. Among FDA pregnancy category D and X compounds, diclofenac and atorvastatin were negative in the rat embryonic-fetal toxicity study, but positive in the ZET and mEST. FDA pregnancy category D and X compounds, except thalidomide and ribavirin, were judged as positive in the ZET or mEST. These results suggest that the combination of the ZET and mEST is a good prediction model for developmental toxicity. Conversely, ribavirin was judged as negative in both the ZET and mEST, but positive in the rat embryonic-fetal toxicity study. It is unclear why the ZET and mEST could not detect the toxicity of ribavirin at present. Comparing the results between the ZET and mEST, they identified different types of true positives and true negatives (Figs. 2 and 3). We believe better prediction results would be achieved using both the ZET and mEST.

In summary, we found that the ZET and mEST are adequate assessment models for developmental toxicity screening in a high-throughput manner. Complementary use of the two models should improve the prediction

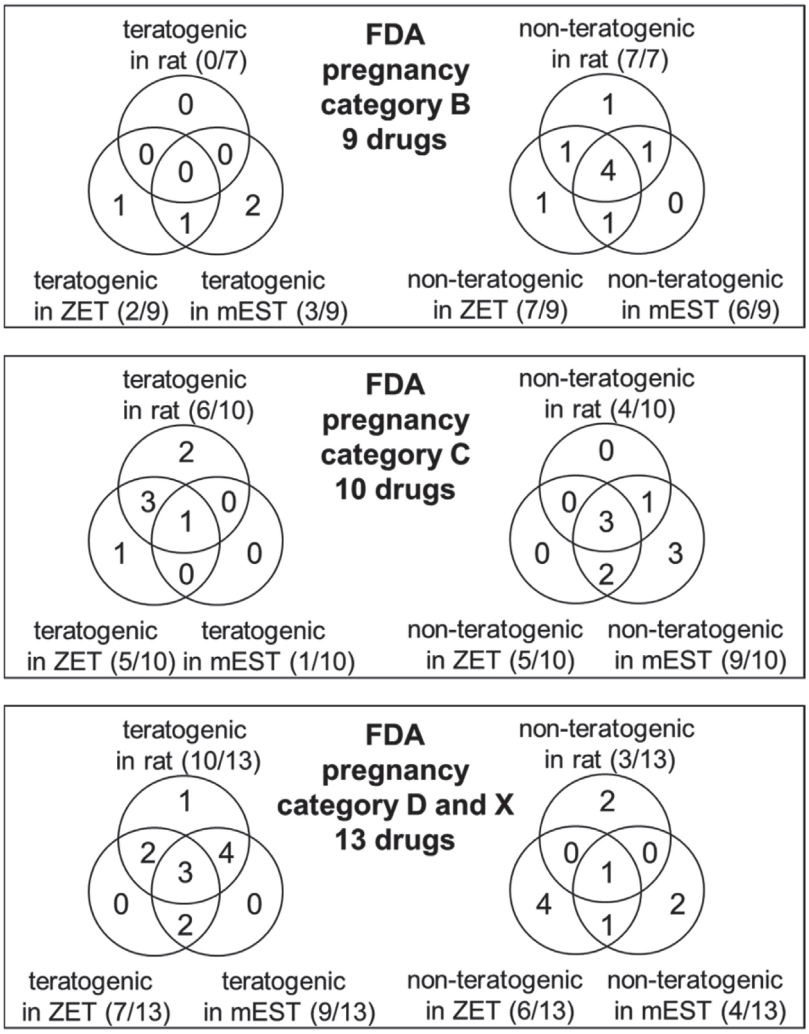

Fig. 4. Venn diagrams of the results from the ZET, mEST, and rat embryonic-fetal toxicity study using 31 compounds including human pharmaceutical drugs.

accuracy of the teratogenicity of drug substances. These two models would contribute not only to the evaluation of potential developmental toxicity at the early stage in drug development but also to decreasing the use of experimental animals compared with rat embryonic-fetal toxicity studies. Additionally, the two models would lead to the elucidation of structure and activity relationships that are important for pharmaceutical companies. Thus, these models would greatly improve identification of low developmentally toxic compounds in future drug discovery programs.

ACKNOWLEDGMENTS

We thank Reiko Kawase, Junko Koiwa, Hiroko Nakayama, Yuka Hayakawa, Yuka Takahashi, Chizuru Hirota, and Michiko Ariyoshi for assistance with the experiments, and Rie Ikeyama and Yuka Mizutani for secretarial support.

Conflict of interest---- AI, NM, TM, MM, JK, TS, YH, and SK are employees of Ono Pharmaceutical Co., Ltd (Osaka, Japan). KK was an employee of Ono Pharmaceutical Co., Ltd. YN, NU, YS, and TT received a research grant from Ono Pharmaceutical Co., Ltd.

REFERENCES

Aikawa, N., Kunisato, A., Nagao, K., Kusaka, H., Takaba, K. and Ohgami, K. (2014): Detection of thalidomide embryotoxicity by *in vitro* embryotoxicity testing based on human iPS cells. J. Pharmacol. Sci., **124**, 201-207.

Brannen, K.C., Panzica-Kelly, J.M., Danberry, T.L. and 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.

De Jong, E., van Beek, L. and Piersma, A.H. (2014): Comparison of osteoblast and cardiomyocyte differentiation in the embryonic stem cell test for predicting embryotoxicity *in vivo*. Reprod Toxicol., **48**, 62-71.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N., Piersma, A., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., Bremer, S. and 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. Altern. Lab. Anim., **30**, 151-176.

Hagihara, K., Nishiya, Y., Kurihara, A., Kazui, M., Farid, N.A. and Ikeda, T. (2008): Comparison of Human Cytochrome P450 Inhibition by the Thienopyridines Prasugrel, Clopidogrel, and Ticlopidine. Drug Metab. Pharmacokinet., **23**, 412-420.

ICH (2005) : Detection of Toxicity to Reproduction for Medicinal Products & Toxicity to Male Fertility S5 (R2).

ICH (2015) : Final Concept Paper S5 (R3): Detection of Toxicity to Reproduction for Medicinal Products & Toxicity to Male Fertility.

Ito, T., Ando, H., Suzuki, T., Ogura, T., Hotta, K., Imamura, Y., Yamaguchi, Y. and Handa, H. (2010): Identification of a primary target of thalidomide teratogenicity. Science, **327**, 1345-1350.

Matsumoto, N. (2009): Prediction of Embryotoxicity by Embryonic Stem Cell Test. AATEX. **14 (supp.)**, 1018.

Mosmann, T. (1983) : Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J. Immunol. Methods, **16**, 55-63.

Nishimura, Y., Inoue, A., Sasagawa, S., Koiwa, J., Kawaguchi, K., Kawase, R., Maruyama, T., Kim, S. and Tanaka, T. (2016): Using zebrafish in systems toxicology for developmental toxicity testing. Congenit. Anom. (Kyoto), **56**, 18-27.

OECD (2013) : OECD GUIDELINES FOR THE TESTING OF CHEMICALS, Fish Embryo Acute Toxicity (FET) Test.

Riebeling, C., Fischer, K., Luch, A. and Seiler, A.E. (2015): Classification of reproductive toxicants with diverse mechanisms in the embryonic stem cell test. J. Toxicol. Sci., **40**, 809-815.

Yamashita, A., Inada, H., Chihara, K., Yamada, T., Deguchi, J. and Funabashi, H. (2014): Improvement of the evaluation method for teratogenicity using zebrafish embryos. J. Toxicol. Sci., **39**, 453-464.