

*IN VITRO* MORPHOGENESIS IN DEVELOPMENTAL TOXICITY TESTING:  
ASSAY VALIDATION AND THE SEARCH FOR TERATOGENIC MECHANISMS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAII AT  
MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

DEVELOPMENTAL AND REPRODUCTIVE BIOLOGY

MAY 2018

By  
Erica L.L. Warkus

Dissertation Committee:

Yusuke Marikawa (Chairperson)  
W. Steve Ward  
Vernadeth Alarcon  
Takashi Matsui  
Joe Ramos

Key words: Erica Warkus, Developmental Toxicity, *In Vitro* Gastrulation, Embryoid Body Morphogenesis,  
Validation of Stem Cell Models, Fluoxetine

ProQuest Number: 10992921

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10992921

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

## DEDICATION:

*To Scrappy:*

*(Nov. 23, 2000 – Oct. 6, 2017)*

*You danced when I laughed.  
You dried my tears when I cried.*

*You understood me better than anyone on  
the face of this earth, and you boldly lived  
up to your name.*

*You were fearless, resilient and  
unashamedly true to yourself.*

*You taught me how to love and cherish  
myself and others. I will always strive to be  
worthy of your unrelenting devotion.*

*I love you.*

## ACKNOWLEDGEMENTS

After three years, I am finally writing this note of gratitude as the finishing touch to my dissertation. This has been a time of personal growth, intensive learning and many life changes. I would like to acknowledge the people who have helped and supported me for the duration of this project.

I would first like to thank my mentor, Dr. Yusuke Marikawa. When I started this journey with you, I was often obstinate with a habit of charging off on tangents. Somehow you kept me on track and channeled all that energy into the eventual creation of this very dissertation, all the while letting me think it was my idea. I will never fully understand how you did it. I think you must be omniscient, because you always seem to know all the lab supplies I'll need or the experiments I'll do long before I'm aware of either. You also always seem to be right—it's rather impressive. I admire you immensely and will always be deeply grateful to have spent these years learning from your wisdom and guidance.

I would also like to thank the people that helped me have a semblance of a social life during the last three years. Brent, you are the main reason that I surf, run, listen to podcasts, practice jujitsu or have intense philosophical discussions in the culture room or lab at 11 p.m. Trevor, you were always happy to meet up for a beer and a chat. I admire your reliability and the way that you just get things done. Aileen, your patience is endless and you have a quiet but sharp humor that makes me laugh everytime. Charlie, you helped me stay centered and showed me how to find myself. Kai, dancing with you is the perfect cure for a tired brain. To Jaeger, Evan, Elliott, Amy, Britt, Nisha, Michael, Mattie, Sam, Emily (congresswoman), Emily (dreads), Megan, Shaun, Hero, Phil, Ross, Mike, Jen, Stefan, Sunny, Tim, Ron and Bianca: Don't change a thing. I love you all.

I would also like to thank my myriad parents for the love and support that you gave me. Terry, you taught me how to fix my car and build a computer (now two of my favorite hobbies). Mom, you gave me your positivity and creative energy. Momma, you taught me how to be quiet. Mike, you taught me how solve problems rationally and how to listen to people's stories as well as to write my own. Auntie Lissa, you taught me grace and you always made sure I had a welcoming family to come home to on Oahu.

Thank you from the bottom of my heart.

Sincerely,

Erica Lyn Lam Warkus

March 29, 2018.

## ABSTRACT

Teratogens are chemicals that can cause birth defects in the developing embryo during pregnancy. Major structural birth defects affect 2-5% of children, but the etiology of this developmental error is often unknown. Exposure to teratogens (chemicals that cause birth defects) during gestation may play a causative role in many of these congenital defects. Therefore, it is important to identify teratogenic chemicals before pregnant women are exposed to them, but there is no easy, reliable test to determine potential teratogenicity of pharmaceutical drugs or industrial chemicals (pesticides, herbicides, solvents, etc.). Current developmental and reproductive toxicity (DART) testing uses millions of pregnant animals in outdated, inefficient regulatory studies. Embryonic stem cell-based *in vitro* tests (EST) produce faster, clearer results and have the potential to revolutionize DART testing, but they often lack the biologic complexity created during embryonic gastrulation and morphogenesis. The embryo is most sensitive to teratogenic exposures during the first three to eight weeks of fetal life. During this time, the process of morphogenesis shapes nearly all three-dimensional (3-D) organs and tissue structures in the embryo. Many of the common birth defects (e.g., neural tube closure defects and cardiac septal defects) are caused by disruptions in this vital process. Since morphogenesis drives the formation of fetal structures and is often affected by teratogenic drugs, a screen for potential teratogens should incorporate similar 3-D structural complexity, but this is a common deficit of *in vitro* teratogenicity tests. Aggregated P19C5 stem cells spontaneously form “embryoid bodies” (EBs) and recapitulate the germ layer differentiation and structural remodeling that occur during embryonic gastrulation and morphogenesis. Exposure to teratogens disrupts EB morphology, making P19C5 morphogenesis a sensitive indicator of developmental toxicity. In this research project, I validated the P19C5 system as a morphology-based teratogen-screening assay and demonstrate that the P19C5 assay can be used to investigate teratogenic mechanisms that cause structural defects. We expect that this test will allow us to accurately identify chemicals that may cause fetal malformations and help us understand the mechanisms of teratogenic chemicals in order to reduce the number of children born with preventable birth defects.

# CONTENTS

Acknowledgements .....	iii
Abstract .....	iv
List of Tables .....	vii
List of Figures .....	viii
List of Abbreviations .....	ix
Chapter 1. Background .....	1
1.1 - Birth defects and teratogens .....	1
1.2 - Identification of teratogens: Human epidemiology .....	2
1.3 - Identification of teratogens: Animal studies .....	4
1.4 - Identification of teratogens: <i>In vitro</i> models .....	6
1.5 - A new tool for teratogen research: The P19C5 <i>in vitro</i> gastrulation model .....	8
Chapter 2. Validation of the P19C5 Morphogenesis Model – Category X drugs .....	13
2.1 - Introduction .....	13
2.2 - Materials and methods .....	15
2.2.1 Cell culture and embryoid body formation .....	15
2.2.2 Drugs .....	16
2.2.3 Cell viability assay .....	16
2.2.4 Morphometric parameters .....	16
2.2.5 Statistics .....	17
2.2.6 Experimental design .....	17
2.3 - Results .....	19
2.3.1 Time course of growth and axial elongation morphogenesis in P19C5 EBs .....	19
2.3.2 Drugs that impacted P19C5 cells with high potency .....	19
2.3.3 Drugs that impacted P19C5 cells with mild potency .....	20
2.3.4 Drugs that did not exhibit any impact .....	21
2.3.5 Correlation between drug impact on P19C5 cells and known pregnancy risks .....	22
2.4 - Discussion .....	23
Chapter 3. Exposure-Based Validation of an <i>In vitro</i> Gastrulation Model for Developmental Toxicity Assays .....	38
3.1 - Introduction .....	38
3.2 - Materials and methods .....	40
3.2.1 Test compounds .....	40
3.2.2 Cell culture .....	40
3.2.3 Image analysis .....	40
3.2.4 Viability assay .....	41
3.2.5 Statistical analyses .....	41
3.3 - Results .....	43

3.3.1 Experimental design.....	43
3.3.2 Effects of the positive and negative exposures on EB morphogenesis.....	43
3.3.3 Cytotoxicity on somatic cell lines .....	44
3.4 - Discussion.....	46
Chapter 4. Fluoxetine inhibits canonical Wnt signaling to impair embryoid body morphogenesis: potential teratogenic mechanisms of a commonly used antidepressant .....	59
4.1 - Introduction .....	59
4.2 - Materials and methods .....	62
4.2.1 Compounds.....	62
4.2.2 Cell culture and embryoid body generation .....	62
4.2.3 Morphometric analyses .....	62
4.2.4 Reverse transcription and polymerase chain reaction (RT-PCR) assays .....	63
4.2.5 Luciferase reporter assays.....	63
4.2.6 Cell viability assay.....	64
4.2.7 Statistics.....	64
4.3 - Results .....	65
4.3.1 Fluoxetine impairs morphogenesis of P19C5 embryoid bodies in a dose-dependent manner .....	65
4.3.2 The morphogenetic effects of fluoxetine is not due to inhibition of serotonin transporter .....	65
4.3.3 Fluoxetine alters expression patterns of various developmental regulator genes.....	66
4.3.4 Fluoxetine inhibits canonical Wnt signaling in P19C5 cells .....	67
4.3.5 Activation of canonical Wnt signaling partially alleviates the adverse effects of fluoxetine .....	68
4.3.6 Trifluoromethylphenyl moiety of fluoxetine is essential in causing the adverse effects of fluoxetine .....	69
4.3.7 Fluoxetine diminishes cell proliferation independently of its inhibitory effects of Wnt signaling or SERT .....	69
4.4 - Discussion.....	71
Chapter 5. Conclusion.....	86
5.1 - The future of developmental toxicology research is <i>in vitro</i> .....	86
5.2 - The P19C5 morphogenesis model adds complexity to <i>in vitro</i> developmental toxicity screens.....	89
5.3 - Future directions and concluding remarks.....	90
Appendix A: Macro Script for ImageJ morphometric analysis of EBs .....	92
Appendix B: Example images corresponding to Appendix A macro script for ImageJ morphometric analysis of EBs.....	95
References.....	97

# LIST OF TABLES

## CHAPTER 2

Table 2.1. Compounds used in the Category X validation study .....	37
--	----

## CHAPTER 3

Table 3.1: Compounds used in the present study Daston Validation study .....	56
Table 3.2: Adverse impacts of the Daston positive exposures on EB morphology .....	57
Table 3.3: Adverse impacts of the Daston negative exposures on EB morphology .....	58

## CHAPTER 4

Table 4.1: Genes examined in the fluoxetine study .....	84
Table 4.2: Compounds examined in the fluoxetine study .....	85



# LIST OF FIGURES

## CHAPTER 1

Figure 1.1: Diagram depicting teratogenic susceptibility during development .....	12
---	----

## CHAPTER 2

Figure 2.1: Temporal changes in EB morphology and experimental schema .....	28
Figure 2.2: Drugs that impacted P19C5 cells with high potency .....	29
Figure 2.3: Drugs that impacted P19C5 cells with mild potency .....	30
Figure 2.4: Drugs that did not impact P19C5 cells .....	31
Figure 2.5: Summary diagram of drug impacts .....	32
Figure 2.6: Representative images of P19C5 EBs .....	33-36

## CHAPTER 3

Figure 3.1: Experimental schema for the Daston Validation study .....	50
Figure 3.2: Impact of the Daston compounds on P19C5 EB morphogenesis .....	51-52
Figure 3.3: Representative images of P19C5 EBs .....	53
Figure 3.4: Cellular proliferation and viability assay .....	54

## CHAPTER 4

Figure 4.1: Experimental schema for investigating the mechanism of fluoxetine .....	75
Figure 4.2: Fluoxetine impairs morphogenesis of P19C5 embryoid bodies .....	76
Figure 4.3: Morphogenetic effects of fluoxetine are not due to inhibition of the serotonin transporter .....	77
Figure 4.4: Fluoxetine alters gene expression patterns of developmental regulator genes .....	78
Figure 4.5: Fluoxetine inhibits canonical Wnt signaling in P19C5 cells .....	79
Figure 4.6: Activation of canonical Wnt signaling partially alleviates the adverse effects of fluoxetine .....	80
Figure 4.7: Trifluoromethylphenyl moiety of fluoxetine is essential in causing adverse effects .....	81
Figure 4.8: Fluoxetine decreases cell proliferation independently of effects on Wnt or SERT activity .....	82
Figure 4.9: Proposed mechanism of action by which fluoxetine affects morphogenesis .....	83

## LIST OF ABBREVIATIONS

3-D	Three-dimensional	mESC	Mouse embryonic stem cells
5-FU	5-fluorouracil	MIE	Molecular initiating event
5-HT	Serotonin (5-hydroxy tryptamine)	mRNA	Messenger RNA
AOP	Adverse outcome pathway	NA	Not aggregated
ASD	Atrial septal defect	NET	Norepinephrine transporter
ATP	Adenine triphosphate	NF	Norfluoxetine
ATRA	All-trans retinoic acid	NHP	Non-human primate
AUC	Area under the curve	NIH	National Institute of Health
AVSD	Atrioventricular septal defect	NOAEL	No observable adverse effect level
BPA	Bisphenol A	NPC	Neural progenitor cells
CAS/CAS RN	Chemical Abstracts Service Registration Number	NRI	Norepinephrine reuptake inhibitor
Cmax	Maximum plasma concentration	NTD	Neural tube defect
CTNNB1	$\beta$ -catenin	OECD	Organization for Economic Co-operation and Development
DART	Developmental and reproductive toxicity	RA	Retinoic acid
DDT	Dichlorodiphenyltrichloroethane	RAR	Retinoic acid receptor
DES	Diethylstilbestrol	RARE	Retinoic acid receptor element
DMSO	Dimethyl sulfoxide	RF	R-fluoxetine
DNA	Deoxyribonucleic acid	ROS	Reactive oxygen species
EB	Embryoid body	RT-PCR	Reverse transcriptase polymerase chain reaction
ECVAM	European Centre for the Validation of Alternative Methods	RVOTO	Right ventricular outflow tract obstruction
EDI	Elongation distortion index	RXR	Retinoid X receptor
EMT	Epithelial to mesenchymal transition	SERT	Serotonin transporter
ER	Endoplasmic reticulum	SEURAT	Safety Evaluation Ultimately Replacing Animal Testing
ES/ESC	Embryonic stem cells	SF	S-fluoxetine
EST	Embryonic stem cell test	Shh	Sonic hedgehog
ESTc	Cardiomyocyte EST	SNRI	Serotonin norepinephrine reuptake inhibitor
FAS	Fetal alcohol syndrome	SSRI	Selective serotonin reuptake inhibitor
FASD	Fetal alcohol spectrum disorder	TDS	Testicular dysgenesis syndrome
FDA	Food and Drug Administration	TFMP	Trifluoromethylphenyl
GI	Gastrointestinal	TNKS	Tankyrase
hESC	Human embryonic stem cells	TT21C	Toxicity Testing in the 21st Century
HMG-CoA	$\beta$ -Hydroxy $\beta$ -methylglutaryl-CoA	v/v	Volume/volume
Hox	Homeobox	VPA	Valproic acid
LHRH	Luteinizing hormone-releasing hormone	VSD	Ventricular septal defect
LOAEL	Lowest observable adverse effect level	WE/WE	Whole embryo/whole embryo culture
MAPK	Mitogen-activated protein kinase	WMISH	Whole-mount <i>in situ</i> hybridization
MEHP	Mono(2-ethylhexyl) phthalate		

# CHAPTER 1. BACKGROUND

## 1.1 - Birth defects and teratogens

Major structural birth defects affect 3% of newborns, but the etiology of the congenital defect is unknown in nearly 50% of these cases (Rynn *et al.*, 2008). Environmental exposures, including exposure to teratogenic chemicals may play a causative role in many of the unknown cases (Sadler, 2012). Birth defects are the result of abnormal development due to either genetic mutations or external factors that disrupt normal developmental processes in specific tissues or structures without causing overt embryonic or maternal death. These external factors are called teratogens, a broad term that includes drugs and chemicals, radiation, viral or bacterial infections, hyperthermia, and maternal metabolic conditions. A culturally relevant example of a viral teratogen is the Zika virus, which crosses the placenta and preferentially infects and kills the neural progenitor cells of the developing embryos leading to microcephaly in affected infants (Russo *et al.*, 2017). Developmental toxicants (teratogenic chemicals) are by far the largest class of teratogens, and the number of chemical teratogens that we are exposed to on a daily basis continually increases as we synthesize and process new compounds and materials. Over the course of her pregnancy, a woman may encounter many developmental toxicants in the form of medications, supplements, occupational exposures, contaminated food or water, secondhand smoke, alcohol and illicit drugs (Jacqz-Aigrain and Koren, 2005).

The concept of teratogens was not widely accepted until the early 1960's. Previously, the scientific and medical community assumed that most birth defects were caused by genetic mutations and that the embryo was fully protected from any harmful environmental factors by the placenta. We now know, of course, that both natural and man-made teratogens have been causing birth defects for a very long time. For example, cyclopamine and jervine, two naturally occurring chemicals produced by the California corn lily (*Veratrum californicum*), cause holoprosencephaly or cyclopia in offspring of animals that graze on it. These two teratogens inhibit cholesterol synthesis and the Sonic hedgehog (Shh) signaling pathway. Since Shh is a crucial regulator of facial morphogenesis, its inhibition leads to cyclopia (Gilbert, 2010). A similar phenotype can be created in humans or animals with genetic mutations in the Shh pathway, demonstrating that very similar syndromes of birth defects can be caused by genetic or teratogenic factors that both disrupt the same developmental process.

There are also some naturally occurring compounds that are necessary for normal development, but that become teratogenic if the embryo is exposed to excessively high or low levels. The best example of this is retinoic acid (RA), the active form of vitamin A. Like Shh, the RA signaling pathway is another of the major signaling pathways that regulate normal embryonic development (Niederreither *et al.*, 1997; Tonk *et al.*, 2015; Sakai *et al.*, 2001; Rhinn and Dolle, 2012). In its active form, RA regulates the spatial and temporal patterns of homeobox (Hox) gene expression as well as interacting with the other major

signaling pathways (i.e., the Wnt, Fgf, Notch, Nodal and Bmp pathways) to establish the anterior-posterior axis and the embryonic body plan early in development. Because RA is responsible for patterning and development of the pharyngeal arch derivatives and the caudal body, abnormal levels of RA (or chemicals that disrupt RA signaling) are highly likely to cause characteristic defects that include small or absent ears and jaws, cleft palate, aortic arch abnormalities and neural tube defects (NTD) (Tonk *et al.*, 2015). This demonstrates that some teratogens cause predictable birth defects based on their molecular mechanism and the developmental process that they disrupt.

Endocrine disruptors are another class of teratogenic compounds that are increasingly ubiquitous in our society. The endocrine system is composed of glands that secrete endocrine, paracrine and autocrine hormones as well as the receptors that respond to these secreted hormones. An endocrine disruptor is defined as *“an exogenous agent that interferes with the production, release, transport, metabolism, binding, action or elimination of natural hormones in the body responsible for the maintenance of homeostasis and the regulation of developmental processes”* (Kavlock *et al.*, 2012). The development of embryonic genitalia is particularly sensitive to disruption by variability in the levels of the steroid sex hormones (e.g., androgens and estrogens). Therefore, chemicals that interfere with normal androgen or estrogen signaling have potent effects in embryos, even at levels that are innocuous in adult humans. A notorious example is bisphenol A (BPA), which is used as a stabilizer in polycarbonate plastics and is produced at a rate of 2 billion pounds per year in the US. Unfortunately, BPA leeches out of plastics into water or food at levels that are high enough to alter the reproductive organs of developing frogs, mice, primates and humans. In humans, exposure to BPA in utero causes a prostate enlargement and a decline in sperm count in men and increases the rate of mammary tumors in women. Other endocrine disruptors include dichlorodiphenyltrichloroethane (DDT; a now-banned pesticide), phthalates, acrylamide, nonylphenol and genistein, to name a few. It is suspected that estrogenic endocrine disruptors are responsible for the increasing rates of “testicular dysgenesis syndrome” (TDS) in human males worldwide. Male infants born with TDS have poorly formed external genitalia, low sperm counts and a higher risk of testicular cancer (Gilbert, 2010).

## **1.2 - Identification of teratogens: Human epidemiology**

Clearly it is important to identify potential teratogens and prevent maternal exposures to teratogenic chemicals. Unfortunately, it is difficult to study teratogens in mammalian pregnancies and the current methods used to screen for teratogens—retrospective epidemiologic studies on humans and *in vivo* testing using animal models—are ineffective and inefficient, respectively.

Although retrospective epidemiologic studies on birth defects in human pregnancies have identified a few of the earliest and most notorious teratogens, they are an ineffective way of preventing the burden of teratogen-induced birth defects in human pregnancies. To achieve a sample size sufficient

to identify a drug or medication as a teratogen in an epidemiologic study, the teratogenic agent in question must have already caused thousands of birth defects that could easily have been prevented. For instance, thalidomide was widely prescribed to pregnant women in the 1960's to treat morning sickness and was only identified as a teratogen once the medical field noted a high incidence of babies born without arms or legs, a condition called phocomelia (Gilbert, 2010).

Human gestation is divided into two stages: the embryonic period in the first eight weeks and the fetal period spanning the remainder of gestation. Most teratogen-induced birth defects occur between three and eight weeks after conception. In the first three weeks of the embryonic period, teratogens are unlikely to cause congenital malformations because the embryo has not undergone gastrulation and the formation of the three germ layers. Before embryonic cells become committed to a single cell lineage, they are pluripotent, meaning they can become any of the cell types in body. At this stage, a teratogenic insult is likely to either kill the entire embryo or damage only a few pluripotent cells without causing lasting effects in the embryo. At approximately three weeks post-conception, the pluripotent cells of the epiblast become either embryonic ectoderm (the source of neural crest cells, central nervous system [CNS], and surface ectoderm [skin/hair/teeth, etc.]), endoderm (the future epithelial and secretory lining of the lungs and gastrointestinal (GI) tract) or mesoderm (the source of musculoskeletal, cardiovascular and other visceral tissue) through the process, called gastrulation. After gastrulation is the period of organogenesis, during which the pluripotent embryonic cells become committed to specific progenitor cell lineages, each equipped to eventually differentiate into a single organ or tissue type. A teratogenic disruption to a specific progenitor cell population inhibits either the migration, differentiation or survival and proliferation of the cells and has lasting effects that are amplified as the organ or structure continues to develop.

During the embryonic stage, organs and body structures develop at different times and have correspondingly different periods of susceptibility to teratogenic disruption. Thalidomide causes a specific sequence of abnormalities that depend on the timing of the embryo's exposure (Fig. 1.1A). Although it is only teratogenic between 20 to 36 days of gestation, a single dose of thalidomide can cause a baby to be born with all four limbs absent or severely shortened. If the exposure occurs in day 20 to day 24, only the ears are abnormal. Exposure between day 24 to day 32 results in phocomelia of the arms and legs, whereas exposure on day 32 to 36 only causes abnormalities in the formation of the thumbs. This demonstrates an important point, teratogenic effects are determined by the timing, dose and duration of the teratogenic exposure. Despite the dramatic malformations that thalidomide causes, nearly 10,000 children were born with thalidomide syndrome before the connection between thalidomide and phocomelia was made and it was withdrawn from the market (Gilbert, 2010). Not all teratogenic syndromes are distinctive enough to attract attention and implicate a single teratogenic agent, and this is a weakness of identifying teratogenic chemicals through epidemiologic studies.

Alcohol is a more subtle developmental toxicant than thalidomide, but it is undeniably one of the most devastating teratogens due to its high prevalence of use and the severity of the debilitation it causes. Fetal alcohol syndrome (FAS) was described in children born to alcoholic mothers in 1968. FAS is characterized by stunted growth and distinctive facial defects that include a small head, thin upper lip, smooth philtrum and small palpebral fissures of the eyes. Children with FAS often also have structural defects in their hearts, kidneys, or skeleton in addition to the characteristic facial features. The defective neuronal development that contributes to the small head size in affected children also results in FAS being the most common congenital mental retardation syndrome, affecting in 1 in 650 children born in the United States. Fetal alcohol spectrum disorder (FASD) is the diagnosis used for children affected by in utero alcohol exposure that do not meet all the criteria for an FAS diagnosis, and FASD is at least three times as prevalent as FAS. Unlike thalidomide, there is wide variability in the type and the severity of alcohol-induced birth defects. Individuals vary widely in their ability to metabolize alcohol and that fact—along with substantial variation in the timing and dosage of fetal exposure—means that FAS and FASD manifest as a broad spectrum of structural defects. Additionally, alcohol appears to cause developmental toxicity via multiple molecular mechanisms, including the disruption of neural crest cell migration and differentiation, interference with cytoskeletal rearrangement, inhibition of normal Shh signaling, impaired cell-cell adhesion and toxicity in the cell populations that form the forebrain, cranial nerves and face.

The variability seen in FAS and FASD is a perfect example of the phenotypic variability that makes it so difficult to use epidemiologic reports to study teratogens in human pregnancies. Also, since the period of highest susceptibility to teratogens very early, a mother may not even be aware that she is pregnant and thus she is unlikely to monitor environmental exposures (Sadler, 2012). Additionally, the noise created by variability in maternal lifestyle, metabolism and individual gene-environment interactions often confounds the effects of heterogeneous environmental exposures. Thus, it is exceptionally difficult to sift through all factors affecting a pregnancy and pinpoint a single causative agent. Epidemiologic studies will continue to be useful, particularly with the conversion to electronic medical records and increasingly comprehensive birth defect reporting systems. However, additional prophylactic screening methods are necessary to prevent unnecessary birth defects in human pregnancies.

### 1.3 - Identification of teratogens: Animal studies

As a result of the thalidomide crisis, the FDA instituted regulations requiring that new medications undergo testing to screen for teratogenic potential. The regulation and study of teratogens has been guided by six principles proposed by James G. Wilson in 1959:

*“Wilson’s Principles:*

- (A) Susceptibility to developmental toxicity depends on the **genotype** of the conceptus and the way this interacts with adverse **environmental** factors.*

- (B) Susceptibility to developmental toxicity varies with the **developmental stage** at the time of exposure to an adverse influence.
- (C) Developmental toxins act in specific ways (**mechanisms**) on developing cells and tissues to initiate sequences of abnormal developmental events (**pathogenesis**).
- (D) The access of adverse influences to developing tissues depends on the nature of the influence (**agent**).
- (E) The four **manifestations** of deviant development are death, malformation, growth retardation, and functional deficit...
- (F) Manifestations of deviant development increase in frequency and degree as dosage increases [**dose-responsiveness**], from the no effect to the totally lethal level.” (Faqi *et al.*, 2012)

The first of Wilson's principles has two parts: 1) certain species respond differently to particular teratogens and 2) individuals of the same species show variation in the frequency and intensity of teratogen-induced abnormalities. As a result, regulatory developmental and reproductive toxicology (DART) testing is done using at least two species of model animals, one rodent and one non-rodent mammal. Additionally, multiple replicates of each experiment are performed to address individual variability within a single species. The second principle highlights the fact that the period of highest teratogenic susceptibility is early in the first trimester, but that different organs have different windows of susceptibility, as seen in the case of thalidomide. In concordance with the second principle, experiments are conducted to assess the effects of chemical exposures at different stages in the pregnancy (i.e., parental exposure pre-conception, and then treatment of pregnant females either early, mid-, or late in gestation). To identify the dose-response curve described in the sixth principle, chemical exposures are administered in a range of concentrations that starts at the “no observable adverse effect level” (NOAEL) and “lowest observed adverse effect level” (LOAEL) to the concentrations at which the chemical begins to cause maternal toxicity. Finally, the endpoints of the regulatory experiments assess whether in utero exposure to the chemical causes: 1) fetal mortality, 2) congenital malformations, 3) alterations to growth, or 4) functional toxicity after birth (principle four).

The testing methods for developmental and reproductive toxicity (DART) have remained essentially unchanged since they were first instituted in 1966. Since early embryonic development in mammals is very similar, animal models are the gold standard method to assess potential developmental toxicity of chemicals within the context of a placental pregnancy. The embryologic process of gastrulation and germ layer formation is orchestrated by a few major developmental signaling pathways—such as Wnt, Bmp, Fgf and retinoic acid—that coordinate the developmental fate of each fetal cell, establishing body axes and directing cellular migration and interactions (Tam and Loebel, 2007; Tam *et al.*, 2006). These crucial signaling pathways are highly conserved across species, often allowing teratogenic effects

to manifest similarly in different species of animal models. Additionally, early mammalian embryos share nearly identical homologous structures (i.e., limbs, heart) and similar body patterning. The specialized adaptations that distinguish adult forms of different species do not appear until structural remodeling occurs later in development. Thus, teratogenic impacts in the first trimester of human pregnancies are often reflected in similar malformations across other non-human species.

However, there are also a number of drawbacks to relying so heavily on *in vivo* teratogenicity assessments. First, there is some variability in the impact of teratogenic chemicals between species. Thalidomide was not initially identified as a teratogen because mice and rats are non-responsive to its teratogenic effects. Although rabbits show slight malformations after in utero exposure, the only reliable *in vivo* DART model for thalidomide is non-human primates (NHP). Additionally, all *in vivo* models require specialized facilities and trained handlers. The high cost of *in vivo* DART testing translates to high costs for drug development and delays the approval process of new therapeutics. It also gives a competitive advantage to the Big Pharma companies because it is prohibitively expensive for smaller companies to bring drugs to market. Third, the extensive experiments required to perform a thorough DART evaluation for a single chemical combined with the large number of compounds that need to be evaluated mean that DART testing has some of the highest rates of live-animal usage in research, raising ethical and animal welfare issues. Finally, it is very difficult to study the molecular mechanisms of teratogenic effects using *in vivo* models—a concept which is addressed in Wilson's third principle, but that is notably absent from regulatory testing of medications and industrial chemicals. Without understanding how a teratogen causes malformations, it is difficult to confidently translate positive (teratogenic) results from animal studies into an accurate prediction of teratogenic risk in humans. Therefore, the need for mechanistic evaluations of teratogenic pathogenesis is at the forefront of developmental toxicity research, but the tools to investigate teratogenic mechanisms are still emerging. Altogether, these issues highlight the need for additional methods in developmental toxicity testing that will refine and complement *in vivo* teratogenicity tests.

## 1.4 - Identification of teratogens: *In vitro* models

*In vitro* testing methods have the potential to augment *in vivo* screens in a way that will increase efficiency and accuracy of DART testing as well as reduce the volume and cost of animal-based tests. *In vitro* methods offer a standardized testing environment that is a simplified version of normal *in vivo* processes. These methods provide a system that is robust to experimental manipulations and is easily accessed for data collection or visualization of results. Another benefit of alternative model systems is the ability to monitor the exact treatment concentrations, whereas in *in vivo* testing, animals are dosed by body weight and the actual fetal exposure level is unclear. Similarly, *in vitro* models provide a platform for experiments that are very difficult or impossible in live animals, including genetic manipulations like transfection, reporter assays, gene knockouts, etc. Additionally, *in vitro* tests may be used in conjunction



with human hepatic enzymes to mimic human metabolism—identifying seemingly harmless “proteratogens” that require enzymatic activation before they cause effects—bypassing the issue of non-human metabolism and individual variability in a way that is nearly impossible using animal-based testing. All of these advantages mean that there has been substantial interest in the development of *in vitro* teratogen screening assays able to easily and accurately identify teratogens. In fact, this was a central focus of the 2017 Teratology Society conference, where leading developmental toxicologists evaluated the current regulatory protocols for developmental toxicity testing and recommended modifications and future directions of the field (Scialli *et al.*, 2018). In the resulting publication, Scialli *et al.* (2018) write that *in vitro* and *in silico* models are the future of teratogenicity testing because they are compatible, “*with hypothesis-driven testing where we take what we know about a compound or close analog and answer specific questions using targeted experimental techniques rather than a one-protocol-fits-all approach*” (Scialli *et al.*, 2018). *In vitro* tests offer the advantage of clearer analytic endpoints and the flexibility to perform mechanistic evaluations to assess how individual chemicals cause teratogenic effects.

Embryonic stem cell (ESC)-based tests have been explored as alternatives to *in vivo* teratogen screening assays for developmental toxicology. ESCs are derived from the inner cell mass of preimplantation embryos. These stem cells are pluripotent, self-renewing and retain the ability to respond to the major signaling pathways that regulate embryonic development. Pluripotency is the capacity to differentiate into any cell type of the embryonic body, including derivatives of all three germ layers (endoderm, mesoderm and ectoderm). ESCs retain the pluripotency of the ICM *in vitro* and can be directed to differentiate into several cell types by altering culture conditions or adding signaling molecules such as growth factors, hormones, chemicals or proteins (Yu and Thomson, 2008). The goal of *in vitro* testing is to provide a detailed and realistic picture that will clarify our understanding of the unclear or unknown processes occurring *in vivo*, and eventually will allow better prediction of teratogenic drug effects in human pregnancies. Thus far, only the cardiomyocyte embryonic stem cell test (ESTc) (Genschow *et al.*, 2002) has been extensively explored by other researchers as a suitable *in vitro* alternative for teratogenicity screening (Estevan *et al.*, 2009). The principle behind the ESTc resulted from Wobus *et al.* observing that when embryonic stem cells (ESCs) are aggregated in hanging drop culture to form embryoid bodies (EBs), the ESCs differentiated into progenitor cells of all three germ layers (Wobus *et al.*, 1988). Treatment with retinoic acid, a well known teratogen, altered the time course and developmental fate of the differentiating ESCs (Estevan *et al.*, 2009). Building on the changes seen in temporal and lineage specification after teratogenic treatment, ESTc uses ESCs that spontaneously differentiate into cardiomyocytes and evaluates a compound’s ability to inhibit both differentiation and proliferation (or viability) of the ESCs (Spiellmann and Liebsch, 2001). After 10 days, ESC shape and contractility are semi-quantitatively measured to determine the extent of teratogenic effects on differentiation. However, the original ESTc is time-consuming, technically challenging, and uses only a single tissue type (cardiomyocytes) for differentiation studies.

Numerous publications have attempted to improve upon the original ESTc design by incorporating additional analytic endpoints, differentiation pathways or quantitative models. However, the variations of EST continue to adhere to the original format, which assesses differentiation and proliferation of a single embryonic cell line. Modifications to EST (additional cell types, molecular endpoints, reporter-based systems, etc.) increased the sensitivity, but also increased the EST's time frame and complexity, making it laborious as a screening test (Estevan *et al.*, 2009). Despite the many variations and modifications, a test based solely on cytotoxicity and cardiomyocyte differentiation may not be comprehensive enough to serve as a broad screening assay for teratogens. According to Trosko and Chang (2010), the “ultimate potential for *in vitro* testing” will be realized using a “3-D *in vitro* micro-environment;” however, the typical EST lacks 3-D structural complexity and other microenvironment components encountered *in vivo* (Trosko and Chang, 2010). This issue was also discussed by Scialli *et al.* (2018) as well, as they wrote, “*Experimental models that reduce a complex biological system to simpler assays have the benefit of facilitating quantitative evaluation of cellular and molecular responses to chemical perturbation but at the drawback of eliminating the cellular interactions and spatial dynamics that make an embryo complex in the first place. When modeling developmental processes and the toxicity that disrupts them, we need to rebuild this complexity.*” (Scialli *et al.*, 2018) Although *in vitro* tests have the potential to greatly reduce the animal-usage burden imposed by teratogenicity testing, only an *in vitro* system that reflects both teratogenic susceptibility and the complexity of early embryonic development can reasonably reduce animal usage in the field of developmental toxicity.

## **1.5 - A new tool for teratogen research: The P19C5 *in vitro* gastrulation model**

The embryo is most sensitive to teratogenic exposures during the period of gastrulation and organogenesis when crucial developmental regulatory pathways—such as Wnt, BMP, Fgf and retinoic acid—coordinate the developmental fate of each fetal cell, establishing body axes and directing cellular migration and interactions (Tam and Loebel, 2007; Tam *et al.*, 2006). These complex events all contribute to embryonic morphogenesis, the process that creates three-dimensional structures and tissues in the embryo. Many of the most common birth defects (neural tube closure defects, cardiac septal defects, cleft palate, etc.) are the result of genetic or environmental disruptions in morphogenetic process (Herion *et al.*, 2014; Wilde *et al.*, 2014). To model *in vivo* effects, a teratogen screening assay should evaluate adverse impacts on differentiation, as well as embryonic body patterning and morphogenesis. Unfortunately, this type of structural complexity is difficult to replicate *in vitro* and is a common deficit in stem cell-based teratogenicity tests (Lee *et al.*, 2012)

To address the challenges faced by former *in vitro* tests, we propose a novel morphogenesis-based stem cell assay as a screen for developmental toxicants. Similarly to the EBs formed by Wobus in

1988, EBs of aggregated P19C5 stem cells undergo a process of gastrulation and germ layer formation (Marikawa *et al.*, 2009; Wobus *et al.*, 1988). Unlike the original ESC EBs or the EST cardiomyocytes, the P19C5 aggregates are able to spontaneously establish and elongate along a cranial-caudal 'body' axis, recreating the process of embryonic morphogenesis *in vitro* (Fig. 2.1A,C). P19C5 EBs appear to recreate the embryonic process of gastrulation and elongation morphogenesis *in vitro*. During gastrulation, the mass of embryonic cells in the epiblast forms three separate germ layers from which all the structures in the body are derived (Tam and Loebel, 2007; Pfister *et al.*, 2007). Gastrulation begins with the formation of the primitive streak, which elongates towards the caudal end of the embryo. The elongation of the primitive streak is driven by epiblast cells that undergo an epithelial-mesenchymal transition (EMT) and migrate through the primitive streak where they become endoderm and mesoderm.

The P19C5 EBs show temporal and spatial gene expression patterns that mirror the gene expression of gastrulating embryos. Many of the key developmental regulator genes known to be expressed in the primitive streak and posterior mesoderm of mouse embryos at E6.5 to E8.5 increase over the course of four days in hanging drop culture, including: *Brachyury*, *Sp5*, *Fgf8*, *Snai1*, *Lhx1*, *Tbx6*, *Hoxb1*, *Cdx2* and Wnt family members (*Wnt3*, *Wnt3a*, *Wnt5a*, and *Wnt8a*). As would be expected in a differentiating cell mass, pluripotency genes (such as *Pou5f1*, *Foxd3* and *Eras1*) decrease after P19C5 EB aggregation (see *Pou5f1* in Fig. 4.4), and the temporal expression patterns follow the sequences seen in gastrulating embryos. Namely, *Brachyury* expression peaks at Day 1 (marking the primitive streak) followed by Day 2 peaks in the expression levels of mesoderm-associated markers, such as *Wnt3a* (EMT) and *Tbx6* (paraxial mesoderm), whereas the expression level of *Meox1* (somitogenesis) does not peak until Day 3 (Marikawa *et al.*, 2009; Lau and Marikawa, 2014; Li and Marikawa, 2015). Additionally, whole-mount *in situ* analysis (WMISH) shows regionalized gene expression in P19C5 EBs that indicates the presences of a cranial-caudal body axis. The consistent, synchronous morphologic change seen in P19C5 EBs as they morph from ovoids on Day 3 to distinctly elongated forms on Day 4 is characteristic of convergent extension, a process uniquely seen in axial and paraxial mesoderm during elongation morphogenesis of the embryo along the cranial-caudal axis (Fig. 2.1A, C).

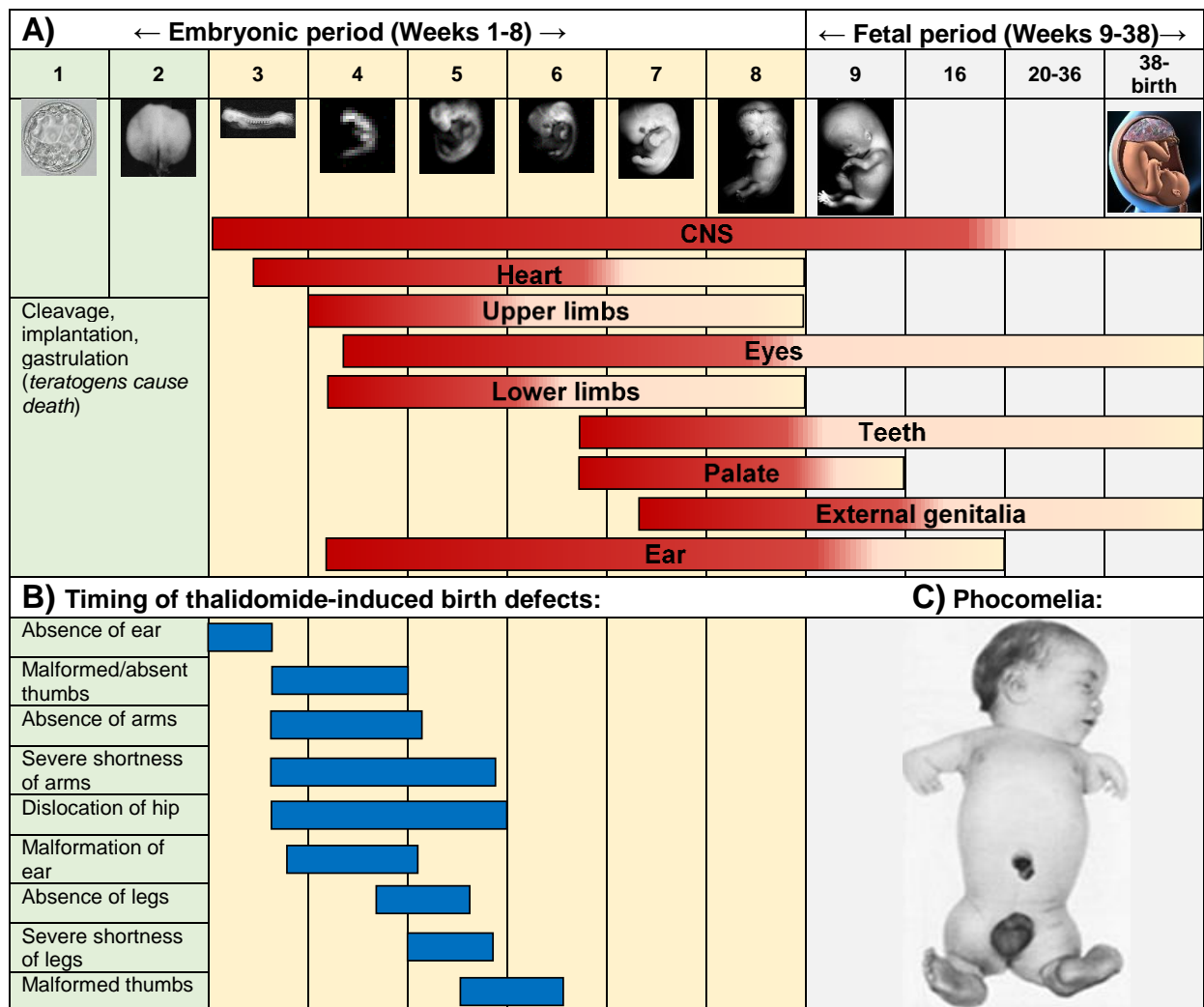
After migratory epiblast cells enter the primitive streak, local embryonic characteristics such as body patterning, cell-cell interactions and 3D structural elements direct the pluripotent cells towards specific development pathways, identifying them as the precursors for a specific tissues or organs (Loebel *et al.*, 2003; Li *et al.*, 2003; Kruegel and Miosge, 2010). This dynamic remodeling of tissue identity and architecture is called morphogenesis. Morphogenesis literally means, "the creation of form," thus, embryonic morphogenesis is aptly named, because it is the process that shapes all the major structural features of the embryonic face, limbs, vital organs, neural tube and body axes during the first three to eight weeks of gestation (Tam *et al.*, 2006; Sadler, 2012). Most of the common birth defects are caused by slight abnormalities in the process of embryonic morphogenesis, and many teratogenic

malformations are the result of exogenous insults to the morphogenetic process (Herion *et al.*, 2014). Experiments using pharmacologic inhibitors of major developmental signaling pathways showed alterations in both P19C5 EB elongation morphogenesis and gene expression profiles like what would be expected if those drugs were used on a developing embryo *in vivo* (Li and Marikawa, 2015). This implies that EB morphogenesis is a sensitive indicator of teratogenic effects and may aid the generation of an *in vitro* alternative to animal-based teratogenicity screens. Additionally, because embryonic morphogenesis depends on the collective functionality of molecular, cellular and tissue-level mechanisms, the parallel process of EB morphogenesis will enhance the analytic power of toxicological models, which attempts to assess dynamic teratogenic mechanisms rather than static analytic endpoints (Gocht *et al.*, 2015; Daston *et al.*, 2015). To the best of our knowledge, this kind of morphogenesis-based *in vitro* assay has never been submitted for validation as a developmental toxicity screening technique.

The P19C5 elongation morphogenesis model has the potential to advance the search for an accurate *in vitro* teratogenicity test, because it encompasses the complexity of gastrulation and germ layer formation in an exceptionally simple, cost-effective hanging drop culture system over four days. The successful elongation of the EBs demonstrates that all the major developmental signaling pathways are intact, allowing quantification of drug impacts using simple image analysis software, ImageJ, which is freely available from NIH. Additionally, the ability to comprehensively investigate the molecular mechanisms of teratogenic effects will allow directed drug development to avoid unintended side effects or off-target actions of new pharmaceuticals. Finally, the simplicity and robustness of the P19C5 system allows for additional modifications such as gene transfection, or human metabolic enzyme additions to more closely mimic the impacts of teratogens on human embryonic development.

The following chapters will describe the work I have done during my Ph.D. project to validate the P19C5 model against two separate panels of chemicals and to use the P19C5 model to understand the previously unknown teratogenic mechanism of a commonly used antidepressant. The first validation study assessed whether the P19C5 model was capable of a binary classification, i.e., teratogenic or non-teratogenic, of 20 drugs known to be either safe (FDA pregnancy Category A or B) or teratogenic (FDA pregnancy Category X) at concentrations of 0.1  $\mu\text{M}$ , 1  $\mu\text{M}$ , 10  $\mu\text{M}$ , and 100  $\mu\text{M}$ . In the Category X validation study, the P19C5 model correctly identified 81% of Category X drugs as potential teratogens using the morphometric parameters of EB area and EDI (elongation distortion index). The Category X validation study was followed by a validation study that tested the P19C5 model against chemicals in the Daston Validation Panel. This study adjusted the concentrations of chemical exposure to match the levels that previously demonstrated teratogenic effects in rodents *in vivo* and, therefore, assessed whether the P19C5 model was capable of responding to teratogens at concentrations comparable to the levels an embryo might encounter during pregnancy. The P19C5 model was able to correctly classify chemicals as teratogenic or safe (within a four-fold concentration margin) with an accuracy of 82%. The third aspect of

my research applied the P19C5 model to a mechanistic investigation of the potential teratogenicity of fluoxetine, a selective serotonin reuptake inhibitor (SSRI) antidepressant. In this portion of my research, I found that fluoxetine appears to inhibit canonical Wnt signaling, an important signaling pathway in development. Finally, in Chapter 5, I will further discuss the implications and applicability of my studies and future research directions.



**Figure 1.1.** Susceptibility to teratogenic disruption is highest during the embryonic period of development and the outcome of a teratogenic exposure depends on the developmental stage at the time of exposure and the nature of the teratogen. (A) Organ systems form at different times and have correspondingly different periods of teratogenic susceptibility. Teratogenic exposure during a susceptible period can cause major structural birth defects (red) or functional defects and minor anomalies (yellow). (B) Exposure to even a single dose of thalidomide causes a syndrome of dramatic congenital malformations. The timing of thalidomide exposure determines the type of malformations seen in affected infants. (C) Thalidomide syndrome was identified in the 1960's due to the unusually high number of exposed infants born with the rare condition, phocomelia, which is characterized by an absence of the long bones of the arms and legs.

## CHAPTER 2. VALIDATION OF THE P19C5 MORPHOGENESIS MODEL – CATEGORY X DRUGS

### 2.1 - Introduction

Exposure to certain chemical agents during pregnancy may disturb the normal course of embryo development. The nature of developmental toxicity varies depending on chemical properties of agents, dosage, and timing of exposure. For example, chemicals that interfere with essential metabolic processes may cause death of early embryonic cells, leading to embryo loss at very early stages of gestation, perhaps even before the pregnancy is detected. Other developmental toxicants may delay embryonic cell cycles, which can result in growth retardation or spontaneous abortion. Cell differentiation may be impaired by chemicals in a tissue type-specific manner, leading to agensis or hypoplasia of the affected organs. Furthermore, developmental toxicants may interfere with cellular behaviors, collectively known as morphogenesis, such as cell migration, adhesion, or shape changes. Morphogenetic disturbance leads to structural anomalies, as found in various common birth defects, such as neural tube closure defects, heart septal defects, hypospadias, and cleft lip and palate (Parker *et al.*, 2010; Sadler, 2012). Because the process of embryogenesis is spatially and temporally complex, integrating cellular proliferation, differentiation, and morphogenesis, identification of developmental toxicants is a highly challenging task and requires multiple investigative approaches.

Many of the chemicals that are considered developmentally toxic or teratogenic to humans were identified through human epidemiologic studies (Friedman, 2009; Chaabane and Bérard, 2013). However, the identification of developmental toxicants using such retrospective methods can be achieved only after a significant number of tragic incidences have already happened. In order to predict potential developmental toxicity of new pharmaceutical agents before they are introduced to the market, chemical regulatory agencies, such as FDA, mandate a series of developmental toxicity tests using model animals (Daston, 2011). However, these animal-based tests also have disadvantages. Metabolic differences between species influence pharmacokinetics of drugs, causing large inter-species variations in the teratogenic dose, possibly resulting in incorrect classifications of potential developmental toxicants (Bailey *et al.*, 2005; Hartung, 2011). Furthermore, regulatory reproductive and developmental toxicology testing is the largest consumer of experimental animals, posing tremendous financial and ethical burdens to the society and the pharmaceutical industry (Hartung and Rovida, 2009). The establishment of effective and economical alternatives to animal-based developmental toxicity tests may ameliorate the impact of metabolic variation between species and reduce the number of animals used in toxicology studies.

*In vitro* toxicity tests using embryonic stem cells, collectively known as embryonic stem cell tests (EST), have been explored as non-animal alternatives in developmental toxicology. Embryonic stem (ES) cell lines derived from the epiblast of pre-implantation stage embryos are capable of differentiating into

various cell types *in vitro* (Yu and Thomson, 2008). In EST, potential developmental toxicity of chemicals is assessed based on their ability to inhibit *in vitro* differentiation of ES cells in relation to their cytotoxic impact on ES cell and somatic cell lines, such as NIH/3T3 fibroblast. The original EST, or ESTc, measures inhibitory effects of chemicals on cardiomyocyte differentiation (Spielmann *et al.*, 1997). ESTc performed well in the initial validation study (Genschow *et al.*, 2002), but it yielded a poor result in a follow-up study (Marx-Stoelting *et al.* 2009; Riebeling *et al.*, 2012). Subsequently, numerous attempts have been made to improve upon the original EST by utilizing additional differentiation pathways (e.g., neurons and osteoblasts) and analytic methods (e.g., expression of tissue-specific molecular markers, transcriptional profiling, and toxicogenomic analysis) (zur Nieden *et al.*, 2001; Buesen *et al.*, 2004, 2009; Seiler *et al.*, 2004; Suzuki *et al.*, 2011; Seiler and Spielmann, 2011; van Dartel and Piersma, 2011; de Jong *et al.*, 2012; Hayess *et al.*, 2013; Panzica-Kelly *et al.* 2013; Gao *et al.*, 2014; Li *et al.*, 2015). However, *in vivo* embryogenesis involves coordinated differentiation and morphogenesis of hundreds of different cell types to construct organized body architectures. Each *in vitro* system represents only a limited aspect of this highly complex process. Therefore, rather than relying on a single system, it would be more realistic and productive to employ multiple *in vitro* models, each of which recapitulates unique aspects of embryogenesis, to detect a wide range of developmental toxicants. Because current versions of EST are based mainly on cell differentiation, additional models incorporating morphogenesis aspects may enhance the versatility of *in vitro* assays and allow the identification of a more diverse array of developmental toxicants.

Previously, we demonstrated that aggregates, or embryoid bodies (EBs), of P19 mouse embryonal carcinoma stem cells exhibit axial elongation morphogenesis *in vitro* (Marikawa *et al.* 2009; Lau and Marikawa, 2014). The P19 cell line was originally isolated from a mouse teratocarcinoma, which was created by transplanting a normal post-implantation embryo into testes (McBurney and Rogers, 1982). P19 cells are developmentally pluripotent, meaning that they are capable of differentiating into derivatives of three germ layers. P19 cells have been used by many researchers as an effective *in vitro* model to investigate the molecular mechanisms of embryonic cell differentiation (McBurney, 1993; Bain *et al.*, 1994, Skerjanc, 1999; van der Heyden and Defize, 2003; Hohjoh, 2013; Voronova *et al.*, 2013). *In vitro* morphogenesis of the P19 subclonal line, P19C5, recapitulates the key aspects of gastrulation driving germ layer formation and convergent extension along the cranial-caudal body axis in early embryos (Marikawa *et al.*, 2009; Lau and Marikawa, 2014). Developmental regulator genes, particularly those involved in gastrulation, mesendoderm formation, axial patterning, caudal extension, and somitogenesis, are expressed during P19C5 EB morphogenesis in a temporally and spatially specific manner comparable to their *in vivo* expression patterns. Furthermore, P19C5 EB elongation is dependent on key developmental signals involved in cell differentiation and embryo patterning, including Wnt signaling pathways (Marikawa *et al.*, 2009), suggesting that the EB morphogenesis is controlled by the same molecular machineries that regulate early development in normal embryos. P19C5 EBs display



consistent size growth and elongation within four days of culture in hanging drops, which are easily accessible for experimental manipulations. Thus, the P19C5 EB model may be used to effectively investigate the impact of chemical agents on embryonic growth, differentiation and morphogenesis.

In the present study, we explored P19C5 EBs as an *in vitro* tool to assess the impacts of 20 therapeutic drugs (Table 2.1). Sixteen of the treatment drugs are contraindicated during pregnancy due to their developmental and reproductive risks (FDA Pregnancy Risk Category X), whereas four are considered safe to be used during pregnancy (FDA Pregnancy Risk Category A). The drugs selected for this exploratory study encompass a wide spectrum of therapeutic targets and chemical properties. The panel of treatment drugs includes agents that affect: 1) lipid metabolism (orlistat [lipase inhibitor] and lovastatin [HMG-CoA reductase inhibitor]); 2) steroidal sex hormone signaling (mifepristone [anti-progesterone], bicalutamide [anti-androgen], diethylstilbestrol [synthetic estrogen], leuprolide [gonadotropin-releasing hormone agonist], and raloxifene [estrogen agonist]); 3) retinoids (acitretin [RAR agonist], and bexarotene [RXR agonist]); 4) monoamine endocrine signaling molecules (doxylamine [anti-histamine], dihydroergotamine [serotonin receptor agonist], metaclopramide [dopamine receptor antagonist]); 5) vitamin metabolism (folic acid, and warfarin [inhibitor of vitamin K epoxide reductase]); 6) antibiotic agents (ribavirin [anti-viral], and nystatin [anti-fungal]); 7) immunomodulatory agents (thalidomide [antiangiogenesis and immunomodulator], and misoprostol [prostaglandin E]); 8) a chemotherapeutic (fluorouracil [antimetabolite]); and 9) an antiarrhythmic (dronedaron [Na<sup>+</sup> channel blocker]). These drugs were also evaluated for their impact on proliferation of P19C5 (embryonic) and NIH/3T3 (somatic) cells to assess general cytotoxicity of the test drugs. The present study shows a strong correlation between the *in vitro* effects and expected developmental toxicity of the drugs, which provides the crucial information on applicability and limitations of the P19C5 EB system as an *in vitro* assay for developmental toxicants.

## **2.2 - Materials and methods**

### **2.2.1 CELL CULTURE AND EMBRYOID BODY FORMATION**

P19C5 cells (Lau and Marikawa, 2014) were maintained in the P19 culture medium (Minimum Essential Medium  $\alpha$ , nucleosides, GlutaMAX Supplement [LifeTechnologies, Carlsbad, CA], 2.5% fetal bovine serum, 7.5% newborn calf serum, 50 units/mL penicillin, and 50  $\mu$ g/mL streptomycin) and passaged every two days when cells had achieved 80-90% confluence in the culture flask. NIH/3T3 cells were obtained from the American Type Culture Collection (Manassas, VA), maintained in the 3T3 culture medium (Dulbecco's Modified Eagle Medium, GlutaMAX Supplement [LifeTechnologies], 10 % fetal bovine serum, 50 units/mL penicillin, and 50  $\mu$ g/mL streptomycin) and passaged every two days. Embryoid bodies (EBs) of P19C5 cells were generated in hanging drops of the culture medium supplemented with 1% dimethyl sulfoxide (DMSO), as previously described (Marikawa *et al.*, 2009; Lau

and Marikawa, 2014). Briefly, P19C5 cells were fully dissociated with Trypsin-EDTA and suspended in the culture medium containing 1% DMSO at a density of 10 cells/ $\mu$ L. Drops (20  $\mu$ L each) of cell suspension were deposited on the inner surface of Petri dish lids for hanging drop culture. All cells and EBs were cultured at 37°C in 4.5% CO<sub>2</sub> in humidified air.

### **2.2.2 DRUGS**

All drugs used in the study were purchased from Sigma-Aldrich (St. Louis, MO; lovastatin [PHR1285], metoclopramide monohydrochloride [M0763], ribavirin [R9644], dronedarone hydrochloride [D9696], leuprolide acetate [L0399], 5-fluorouracil [F6627], orlistat [O4139], doxylamine succinate [D3775], nystatin [N4014] and folic acid [F7876]) or from Santa Cruz Biotechnology (Dallas, TX; acitretin [sc-21075], misoprostol [sc-201264], mifepristone [sc-20126], bicalutamide [sc-202976], diethylstilbestrol [sc-204720], bexarotene [sc-217753], raloxifene hydrochloride [sc-204230], dihydroergotamine methanesulfonate [sc-294343], warfarin sodium [sc-204941] and thalidomide [sc-201445]). All drugs were dissolved in DMSO at 10 mM concentration and stored at -20°C.

### **2.2.3 CELL VIABILITY ASSAY**

The impact of drugs on cell proliferation and viability was evaluated using CellTiter-Glo Luminescent Cell Viability Assay system, which determines the number of live cells in culture by measuring the amount of ATP as a quantitative proxy for the number of metabolically active cells (Promega, Madison, WI). P19C5 cells and NIH/3T3 cells were seeded in 96-well plates at the density of 100 cells/well and 250 cells/well, respectively, and were cultured in the corresponding medium (100  $\mu$ L/well) supplemented with 1% DMSO containing serial dilution of the test drug (0.1, 1, 10, and 100  $\mu$ M) or vehicle only as a control. After 4 days of culture, cells were treated with CellTiter-Glo Reagent for measurement of luminescence, as a readout of ATP amount, according to the manufacturer's instruction (Promega), using Gene Light 55 Luminometer (Microtech, Chiba, Japan). Cell seeding density was optimized through a series of pilot experiments, to confirm that cell numbers at the end of 4 days of culture were proportionate to intensities of luminescence. Relative cell number was calculated based on ratio of the luminescence intensity in drug-treated cells to that in the control of the same set of experiments. For each drug, four sets of experiments were conducted as biological replicates, and the results were presented as mean  $\pm$  standard deviation.

### **2.2.4 MORPHOMETRIC PARAMETERS**

The experimental scheme to test drug impact on EB development is shown in Figure 1D. Each set of experiment consisted of five plates of hanging drops prepared using the same cell suspension: one control plate with no drug and four plates treated with serial dilutions of the test drug (0.1, 1, 10, and 100

μM). Each plate contained 16 hanging drops, which were cultured for four days after initial treatment without medium change. Survival and overall integrity of EBs were monitored on Days 1, 2, and 3. On Day 4, EBs were removed from hanging drops and placed together in a Petri dish filled with phosphate-buffered saline for photography. Images were captured with AxioCam MRm digital camera (Carl Zeiss, Thornwood, NY) attached to Axiovert 200 inverted microscope with Hoffman modulation contrast optics (Carl Zeiss) and controlled by AxioVision software (Carl Zeiss). AxioVision image files were converted to JPG format, which were then opened in ImageJ program (<http://rsb.info.nih.gov/ij>) for morphometric analyses. The circumference of individual EBs was manually traced using the Polygon selections tool to measure the area and circularity (under shape descriptors). In the present study, the area was used to approximate the overall size of EBs. Elongation Distortion Index (EDI) expresses the extent of axial elongation and is calculated as  $(1/[\text{circularity}] - 1)$ , which is equivalent to the formula  $([\text{EB perimeter}]^2/4\pi[\text{EB area}] - 1)$ , as described previously (Marikawa *et al.*, 2009). Area and EDI were normalized by the average values of control EBs in each set of experiments to calculate relative EB size and relative EDI, respectively. Three sets of experiments were conducted for each drug as biological replicates, and all relative EB size and relative EDI were compiled and presented as mean ± standard deviation.

## 2.2.5 STATISTICS

Statistical differences were assessed by two-sample t-test. For cell viability assay, relative cell numbers of P19C5 cells from 4 experiments ( $n = 4$ ) were compared with relative cell numbers of NIH/3T3 cells from 4 experiments ( $n = 4$ ) for each concentration of each drug tested. For EB morphogenesis, relative EDI of drug-treated EBs ( $n = 45$  to 48, compiled from 3 experiments) were compared with relative EDI of control EBs ( $n = 45$  to 48). Differences in average values were deemed significant when  $p$  values were less than 0.01.

## 2.2.6 EXPERIMENTAL DESIGN

Two types of assays were performed to evaluate the impact of the selected 20 drugs at four different concentrations (0.1, 1, 10, and 100 μM). The first assay was to examine drug impact on proliferation of P19C5 and NIH/3T3 cells. Cells were cultured for 4 days as mono-layer at a low density (to minimize contact inhibition) in the presence of test drugs, followed by measurement of total live cell number (see above). Relative cell number of drug-treated cells was normalized by the number of control cells that were cultured for 4 days in the absence of drug. This assay served two purposes: 1) to determine general cytotoxicity and 2) to assess selective inhibition of embryonic cell proliferation. “General cytotoxicity” was defined in the present study as a reduction in relative cell number of NIH/3T3 cells below 20% of control. The rationale for this definition is as follows. Normal NIH/3T3 cells divide

approximately every 20 hours, and therefore they divide an average of 4.8 rounds during 4 days of culture, which increases cell number by  $\approx 27.9$ -fold ( $= 2^{4.8}$ ). If a half of cell divisions are inhibited during this period, cell number increases only by  $\approx 5.3$ -fold ( $= 2^{2.4}$ ), which is approximately 20% or less of control cells ( $= 5.3/27.9$ ). Any drug concentrations that displayed general cytotoxicity were not considered “developmentally toxic” in the present study, regardless of other impacts that are described below. The second purpose of the assay was to assess whether a drug diminishes proliferation of P19C5 cells (representing embryonic cells) more strongly than NIH/3T3 cells (representing somatic cells). This was determined by comparing relative cell numbers between P19C5 and NIH/3T3 cells for each drug treatment. In the present study, significant reduction in P19C5 cell number relative to NIH/3T3 cell number was defined as a sign of selective adverse impact on embryonic cell proliferation.

The second assay examined drug impact on growth and morphogenesis of P19C5 EBs. After 4 days of hanging drop culture in the presence of test drugs, area and EDI were measured as morphometric parameters to assess EB growth and morphogenesis, respectively. In the present study, a reduction in EB size by more than 20% was classified as an adverse drug impact on EB growth. A significant decrease or increase in EDI was classified as an impact on EB morphogenesis. Although mechanistic causes of EDI decreases are likely to be different from EDI increases (see Discussion), both were grouped together as morphogenetic impact in the present study.

Thus, if a drug caused “stronger inhibition of P19C5 proliferation”, “reduction in EB size”, and/or “decrease/increase in EDI” at concentrations that are not “generally toxic”, it was categorized as a developmental toxicant in the present study. Furthermore, if a drug displayed any of these effects at the lowest concentration tested (i.e., 0.1  $\mu$ M), it was considered to have “high potency”. In contrast, the effects observed only at higher concentrations (i.e., 1, 10, and/or 100  $\mu$ M) were considered to have “mild potency”.

## 2.3 - Results

### 2.3.1 TIME COURSE OF GROWTH AND AXIAL ELONGATION MORPHOGENESIS IN P19C5 EBs

During four days of hanging drop culture, P19C5 EBs steadily grew in size, which was quantitatively measured using EB area as a morphometric parameter (Fig. 2.1A, B). The overall shape of EBs was spherical during the first two days of culture (Days 1 and 2) but became ellipsoidal by the third day (Day 3) and distinctly elongated by the fourth day (Day 4; Fig. 2.1A). This temporal change in EB shape was quantified using another morphometric parameter, Elongation Distortion Index (EDI; see Materials and Methods), which significantly increased from Day 2 to Day 3 and from Day 3 to Day 4 (Fig. 2.1C). In the following assessment of drug impact on EBs, the area and EDI were measured on Day 4 to assess overall growth and morphogenesis of EBs.

### 2.3.2 DRUGS THAT IMPACTED P19C5 CELLS WITH HIGH POTENCY

Eight out of the 20 drugs evaluated demonstrated high potency because they impacted P19C5 proliferation and/or EB development at treatment concentrations as low as 0.1  $\mu\text{M}$ . Highly potent drugs included acitretin, bexarotene, diethylstilbestrol, dronedarone, fluorouracil, mifepristone, raloxifene, and ribavirin. Acitretin is a synthetic retinoid that activates retinoic acid receptors and is prescribed for treatment of various skin diseases. Acitretin exhibited general cytotoxicity (i.e., NIH/3T3 cell number < 20%) at 100  $\mu\text{M}$ , which was also reflected in death of P19C5 cells in mono-layer culture as well as in hanging drops by Day 1. At lower concentrations (0.1, 1, and 10  $\mu\text{M}$ ), P19C5 cells aggregated into viable EBs and survived until Day 4, although they showed dramatic decreases in growth and morphogenesis (Fig. 2.2). Acitretin-treated EBs were markedly smaller than control EBs, with an average size less than 50% of control. Additionally, acitretin-treated EBs retained the initial spherical shape void of any discernable elongation, as shown by marked reduction in EDI (Fig. 2.2 & Fig. 2.6A).

Bexarotene, a synthetic retinoid that selectively targets retinoid X-receptors, is used to treat cutaneous T-cell lymphoma. Bexarotene at 100  $\mu\text{M}$  exhibited general cytotoxicity along with death of P19C5 cells in hanging drops by Day 1. EBs were viable for 4 days of culture in other concentrations, but the EB size was markedly reduced. EDI was also significantly reduced by bexarotene at 1 and 10  $\mu\text{M}$  (Fig. 2.2 & Fig. 2.6B).

Diethylstilbestrol (DES) is a non-steroidal anti-estrogen and was historically prescribed to prevent miscarriages and premature deliveries. General cytotoxicity was observed at 100  $\mu\text{M}$ , which corresponded to disintegration of P19C5 EBs by Day 3. EDI was slightly, but significantly increased by DES at 0.1 and 1  $\mu\text{M}$ . DES at 10  $\mu\text{M}$  also reduced the EB size by more than 20% (Fig. 2.2 & Fig. 2.6D).

Dronedarone is a multichannel blocker used to treat cardiac arrhythmia. Dronedarone at 10 and 100  $\mu\text{M}$  exhibited general toxicity and killed P19C5 cells in hanging drops by Day 1. Dronedarone caused significant reduction in P19C5 cell number relative to NIH/3T3 cells at 0.1  $\mu\text{M}$ , but not at 1  $\mu\text{M}$  (Fig. 2.2). No impact on EB development on size or EDI was observed at 0.1  $\mu\text{M}$  or 1  $\mu\text{M}$  (Fig. 2.6G).

Fluorouracil is a pyrimidine analog antimetabolite and an antineoplastic agent used to treat various cancers. It was generally cytotoxic at 10 and 100  $\mu\text{M}$ . At these concentrations of fluorouracil, EBs survived until Day 3 but completely disintegrated by Day 4. Significant reductions in P19C5 cell number and EB size were observed at 1  $\mu\text{M}$ , but not at 0.1  $\mu\text{M}$ . Notably, EDI was significantly increased by fluorouracil at 0.1  $\mu\text{M}$ , and more markedly increased at 1  $\mu\text{M}$  due to the drastically skinny shape of EBs (Fig. 2.3, Fig. 2.6H).

Mifepristone is a synthetic steroid that inhibits progesterone receptor and is used as an abortifacient. It exhibited general toxicity at 100  $\mu\text{M}$ , which also killed P19C5 cells in hanging drops by Day 1. The EB size was significantly reduced by mifepristone at 10  $\mu\text{M}$ . Interestingly, EDI was significantly increased by mifepristone at 0.1  $\mu\text{M}$  but decreased at 10  $\mu\text{M}$  (Fig. 2.2 & Fig. 2.6M).

Raloxifene is a selective estrogen receptor modulator prescribed to prevent osteoporosis in postmenopausal women. General cytotoxicity was observed at 100  $\mu\text{M}$ , which also killed P19C5 cells in hanging drops by Day 1. Similarly to mifepristone, EDI was significantly increased at 0.1  $\mu\text{M}$  but decreased at 10  $\mu\text{M}$  (Fig. 2.2 & Fig. 2.6Q). EB growth was markedly diminished at 10  $\mu\text{M}$  (Fig. 2.2).

Ribavirin is a purine RNA analog used as an antiviral agent. Ribavirin exhibited general cytotoxicity at 100  $\mu\text{M}$ . However, EB growth was only mildly impaired by ribavirin at 100  $\mu\text{M}$  (Fig. 2.2 & Fig. 2.6R). Interestingly, EDI was altered by ribavirin at all concentrations tested, showing a significant increase at 0.1  $\mu\text{M}$ , but decreases at 1, 10, and 100  $\mu\text{M}$  (Fig. 2.2 & Fig. 2.6R).

### **2.3.3 DRUGS THAT IMPACTED P19C5 CELLS WITH MILD POTENCY**

Seven drugs showed mild (i.e., at 1  $\mu\text{M}$  or higher) impacts on P19C5 proliferation and/or EB development, including bicalutamide, dihydroergotamine, doxylamine, lovastatin, misoprostol, nystatin and orlistat. Bicalutamide is a non-steroidal anti-androgen used to treat prostate cancer and hirsutism. General cytotoxicity was observed at 100  $\mu\text{M}$  (Fig. 2.3). At 10  $\mu\text{M}$ , bicalutamide caused significant increase in EDI (Fig. 2.3 & Fig. 2.6C).

Dihydroergotamine is semi-synthetic ergot alkaloid used to treat migraine headaches and orthostatic hypotension. General cytotoxicity was observed at 100  $\mu\text{M}$ , which also killed P19C5 cells in hanging drops by Day 1. Dihydroergotamine at 10  $\mu\text{M}$  concentration reduced both P19C5 cell number

and EB growth but promoted NIH/3T3 cell proliferation. Dihydroergotamine did not show significant impact on EDI (Fig. 2.3 & Fig. 2.6E).

Doxylamine is a first generation anti-histamine used to treat respiratory allergies. At the highest concentration (100  $\mu$ M), doxylamine promoted NIH/3T3 cell proliferation, but it caused a slight reduction in P19C5 cell number. At 100  $\mu$ M concentration, EDI was also significantly decreased, without any observable impact on EB size (Fig. 2.3 & Fig. 2.6E).

Lovastatin is a competitive HMG-CoA reductase inhibitor used to reduce LDL-cholesterol levels in hypercholesterolemia. Lovastatin was generally cytotoxic at 10 and 100  $\mu$ M. Treatment concentrations of 1  $\mu$ M caused a reduction in EB size (< 80% of control). Notably, although no significant change in EDI was observed at 1  $\mu$ M, the EDI standard deviation was markedly larger than the EDI standard deviation of the control (Fig. 2.3), suggesting that lovastatin caused more variability in EB shapes (Fig. 2.6K).

Misoprostol is synthetic prostaglandin E<sub>1</sub> analog used to prevent gastric ulcers induced by non-steroidal anti-inflammatory drugs. It exhibited general cytotoxicity at 100  $\mu$ M. At 10  $\mu$ M, misoprostol caused reductions in P19C5 cell number relative to NIH/3T3 and diminished EB size. Significant increase in EDI was also observed at this concentration (Fig. 2.3 & Fig. 2.6N).

Nystatin binds to cell membrane ergosterol and is used as an antifungal medication. Nystatin did not exhibit general cytotoxicity at any of the concentrations tested. However, a reductions in both EB size and EDI were observed at 100  $\mu$ M (Fig. 2.3 & Fig. 2.6O).

Orlistat is a lipase inhibitor used to treat or prevent obesity. Orlistat reduced P19C5 cell number more significantly than NIH/3T3 at 100  $\mu$ M and caused significant reductions in EB size and EDI (Fig. 2.3 & Fig. 2.6P).

#### **2.3.4 DRUGS THAT DID NOT EXHIBIT ANY IMPACT**

The following 5 drugs did not display general cytotoxicity or show observable impact on P19C5 cells at the concentrations tested: folic acid (vitamin B9, a dietary supplement), leuprolide (a growth hormone-releasing hormone [GnRH] receptor agonist and inhibits gonadotropin secretion from the pituitary gland), metoclopramide (an antagonist of dopamine and serotonin receptors used as an antiemetic), thalidomide (an immunomodulatory agent), and warfarin (a vitamin K epoxide reductase inhibitor used as an anticoagulant) (Fig. 2.4 & Fig. 2.6I, J, L, S, and T).

### **2.3.5 CORRELATION BETWEEN DRUG IMPACT ON P19C5 CELLS AND KNOWN PREGNANCY RISKS**

The overall effects of the 20 drugs evaluated in the present study are summarized in Figure 5. Thirteen of the 16 contraindicated drugs impacted proliferation, EB growth and/or morphogenesis of P19C5 cells, with high or mild potency. Notably, all of the highly potent drugs (acitretin, bexarotene, DES, dronedarone, fluorouracil, mifepristone, raloxifene, and ribavirin) are contraindicated. Of the seven drugs that mildly impacted P19C5 cells, 5 are contraindicated (bicalutamide, dihydroergotamine, lovastatin, misoprostol, and orlistat) while 2 (doxylamine and nystatin) are considered safe for use during pregnancy. The impact of these safe drugs was observed only at the highest concentration (100  $\mu$ M). All of the contraindicated drugs with mild impact, with the exception of orlistat, exhibited effects at concentrations lower than 100  $\mu$ M. Also, although orlistat displayed significant impact only at 100  $\mu$ M, it affected all three parameters of treated P19C5 cells, proliferation, EB size, and EDI. In contrast, doxylamine and nystatin impacted only two aspects parameters. Therefore, the two assays employed in the present study demonstrated a strong correlation between *in vitro* impact and known pregnancy risks of the drugs.



## 2.4 - Discussion

To reduce animal use in developmental toxicology studies, *in vitro* alternatives are strongly desired. Because only a limited aspect of *in vivo* processes can be simulated in a single type of *in vitro* test, multiple tests would be required to encompass a broad range of embryonic events. Here, we explored the potential of P19C5 EBs, which simulate growth and axial elongation morphogenesis of early embryos, as an *in vitro* tool to study human developmental toxicants. The analysis measured morphometric parameters (size and EDI) of Day 4 EBs to evaluate developmental impact of 20 therapeutic drugs with known *in vivo* developmental toxicity in humans (Table 2.1). The morphometric analysis was paired with cytotoxicity data obtained from proliferation assays of embryonic P19C5 cells and of somatic NIH/3T3 fibroblast cells. To our knowledge, this is the first study to incorporate the morphogenesis of EBs as a stem cell-based *in vitro* test for developmental toxicants. Although the effectiveness of P19C5 EB morphogenesis system needs to be further examined using a wider array of chemical compounds with known developmental toxicity, the present study serves as an introduction for a novel *in vitro* test that will complement other non-animal alternative methods.

Recently, another research group has shown that EBs made of mouse ES cells (mESC) also display axial elongation morphogenesis reminiscent of gastrulating embryos (van den Brink *et al.*, 2014). This indicates that *in vitro* recapitulation of morphogenesis is not a property exclusive to P19 and P19C5 cells. An *in vitro* gastrulation model of mESC may be incorporated for developmental toxicity tests in future studies. Nonetheless, elongation morphogenesis appears to occur more easily and prominently in P19C5 EBs compared to mESC EBs. To generate elongating EBs, mESCs are first aggregated in hanging drops for 2 days, before being transferred to a non-adhesive Petri dish filled with a culture medium containing activators of Nodal signaling (Activin A) and Wnt signaling (CHIR99021; a pharmacological inhibitor of glycogen synthase kinase 3) to induce mesendoderm formation. After exposure to these activators, EBs begin to elongate within 2-3 days (van den Brink *et al.*, 2014). In contrast, P19 and P19C5 EBs spontaneously elongate in hanging drops without a change of culture format or addition of Nodal and Wnt activators. The difference between mESC and P19 cells' ability to initiate gastrulation may be due to the difference in the developmental stages represented by these two pluripotent stem cell lines. mESCs retain developmental and molecular properties that are close to the inner cell mass or the epiblast of the pre-implantation embryo, whereas P19 cells possess features of the primitive ectoderm or the epiblast of the post-implantation embryo (Jones-Villeneuve *et al.*, 1982; Yeom *et al.*, 1996; Niwa, 2007). Use of multiple stem cell lines representing different developmental stages is likely to enhance studies of developmental toxicants, because actions of chemical agents may differ depending on embryonic stages. Regardless, the relatively simple protocol to induce morphogenesis in P19C5 EBs offers a practical benefit by reducing the time and labor required to induce gastrulation.

One of the most challenging aspects of developing and validating *in vitro* systems for developmental toxicity tests is the selection of test drugs. Many efforts have been made to choose a suitable list of reference compounds to be used for the validation of *in vitro* as well as non-mammalian test systems. Examples of validation panels include the group of 20 compounds selected under the direction of the European Centre for the Validation of Alternative Methods (ECVAM) (Brown *et al.*, 2002) and the additional 13 compounds chosen for follow-up studies (Marx-Stoelting *et al.*, 2009). Although these compounds were selected carefully based on available, high-quality *in vivo* data, it is controversial as to whether they can serve as “gold standards” for validation of *in vitro* systems (Daston *et al.*, 2010, 2014; Riebeling *et al.*, 2012). In the present study, we selected 20 drugs with diverse chemical and therapeutic properties in order to gain insight into the applicability and limitation of P19C5 EBs to detect developmental toxicants. Although these drugs were selected independently from the previously chosen reference compounds for the other studies, there were some overlaps with the 13 follow-up compounds mentioned above, namely lovastatin, warfarin, doxylamine, and metoclopramide. Nonetheless, it is clear that further investigations are necessary using additional compounds to more fully reveal the strengths and weaknesses of each *in vitro* system.

Because most embryonic events, including cell differentiation and morphogenesis, are tightly regulated by the actions of specific gene products, incorporation of molecular endpoint analysis is likely to allow more effective detection of developmental toxicity. Indeed, various molecular endpoints, particularly involving detection of cell type-specific gene expression, have been incorporated into ESTs, so that chemical impact on cell differentiation can be assessed more sensitively, quantitatively, objectively, and speedily (zur Nieden *et al.*, 2001; Buesen *et al.*, 2004, 2009; Seiler *et al.*, 2004; Suzuki *et al.*, 2011; Seiler and Spielmann, 2011; van Dartel and Piersma, 2011; de Jong *et al.*, 2012; Hayes *et al.*, 2013; Panzica-Kelly *et al.* 2013; Gao *et al.*, 2014; Li *et al.*, 2015). P19C5 EB system is also amenable to the addition of molecular endpoint analyses. Previously, we showed that the Alzheimer’s medication, donepezil, diminishes EB elongation. While the morphogenetic impact of donepezil was detectable in Day 4 EBs as a significant reduction in EDI, the expressions of developmental regulator genes, such as *Fgf8* and *Cdx2*, were significantly altered by donepezil in Day 2 EBs (Lau and Marikawa, 2014). In P19 and P19C5 EBs, a number of developmental regulator genes for gastrulation and axial patterning, such as *Brachyury*, *Snai1*, *Wnt3a*, *Notch1*, and Hox genes, are expressed in a temporally and spatially specific manner (Marikawa *et al.*, 2009; Lau and Marikawa, 2014). Thus, future investigations incorporating gene expression analyses should provide deeper insight into potential developmental toxicity of various chemical agents and their mechanisms of action.

P19C5 EB morphogenesis simulates only a limited stage and a limited region of embryo development. The temporal profile of gene expressions (Marikawa *et al.*, 2009; Lau and Marikawa, 2014) suggests that the four days of P19C5 EB development represent E5.0 to E8.0 of mouse embryonic

stages, which roughly correspond to the second through fourth weeks of human development. Also, based on the spatial profile of gene expressions, P19C5 EBs appear to represent the caudal region of the embryo, but not the cranial structures, such as the brain and the heart. Nonetheless, many of the contraindicated drugs affected P19C5 EB development in the present study. It is possible that numerous developmental regulators that also play essential roles in other stages and regions of embryo development control P19C5 EB morphogenesis. Elongation morphogenesis of P19C5 EBs depends on various core developmental signals, specifically Wnt, Nodal/Activin, Bmp, Fgf, and retinoic acid pathways, which alters EB morphogenesis in a manner consistent with their roles in gastrulation and axial elongation and patterning *in vivo* (Marikawa *et al.*, 2009; unpublished data). Chemicals that interfere with these signals are likely to impair P19C5 EB morphogenesis. Apart from the synthetic retinoids, acitretin and bexarotene, it is currently unknown whether the contraindicated drugs tested in the present study influence the above signaling pathways. However, genetic studies have linked anomalies in these signaling pathways to numerous human birth defects; for example, Wnt signaling is implicated in tetra-amelia and congenital duplication of palm syndrome (Niemann *et al.*, 2004; Al-Qattan *et al.*, 2009), Nodal/Activin signaling is linked to heterotaxy syndrome (Ma *et al.*, 2012), and Fgf signaling is associated with achondroplasia and Apert syndrome (Yamaguchi and Rossant, 1995; Hajihosseini, 2008). Thus, in spite of a limited representation of developmental events, P19C5 EB morphogenesis may be more versatile in detecting a wide range of developmental toxicants.

In the present study, size and EDI were used as morphometric parameters of EBs to quantitatively measure developmental impact of drugs. EDI is calculated based on the size (area) and the length of circumference (perimeter) of individual EBs, and its increase implicates more elongation and/or distortion of the EB shape. EB elongation is driven by convergent extension, a morphogenetic process of orchestrated cell migration and intercalation (Tada and Heisenberg, 2012). Thus, reduction in EDI may be caused by impairment in convergent extension, which is likely to contribute to developmental anomalies *in vivo*, including neural tube defects (Copp *et al.*, 2003; Ueno and Greene, 2003). In contrast, the mechanistic basis for an increase in EDI is currently unclear, although stimulated cell migration or reduced surface tension may cause pronounced elongation or distortion of EBs. Further investigations are required to elucidate the cellular and molecular mechanisms that are behind increased EB elongation or distortion. Also, EDI alone is not sufficient to distinguish between elongation and distortion effectively. Thus, in addition to EDI, other morphometric parameters need to be explored in future studies to detect various morphogenetic impacts on EBs.

Three of the contraindicated drugs (leuprolide, thalidomide and warfarin) did not alter EB size or EDI, whereas two safe drugs (doxylamine and nystatin) impaired EB development at the highest concentration (100  $\mu$ M). These five cases of apparent “misclassifications” may exemplify limitations of the P19C5 model in predicting certain developmental toxicants. Like most other *in vitro* tests, the P19C5 EB

system recapitulates only embryological events, but not maternal environment. Thus, it is unlikely to detect developmental toxicants that primarily act on maternal tissues and indirectly disturb embryo development. This may be the case for leuprolide, which primarily acts on the maternal pituitary gland and alters the levels of endocrine hormone signaling to secondarily impair embryogenesis. Furthermore, most *in vitro* systems, including P19C5 EBs, are limited by a reliance on an artificial culture environment, which may obscure the effects of certain drugs. Developmental toxicity of warfarin is likely due to its action as an inhibitor of vitamin K regeneration. Animal serum in the P19C5 culture medium may contain a substantial amount of vitamin K, which may negate the effects of warfarin. Because most culture media, supplements and sera are enriched with various essential metabolites, results of *in vitro* toxicity tests, particularly those with no impact, may require cautious interpretation with consideration of mechanisms of drug action.

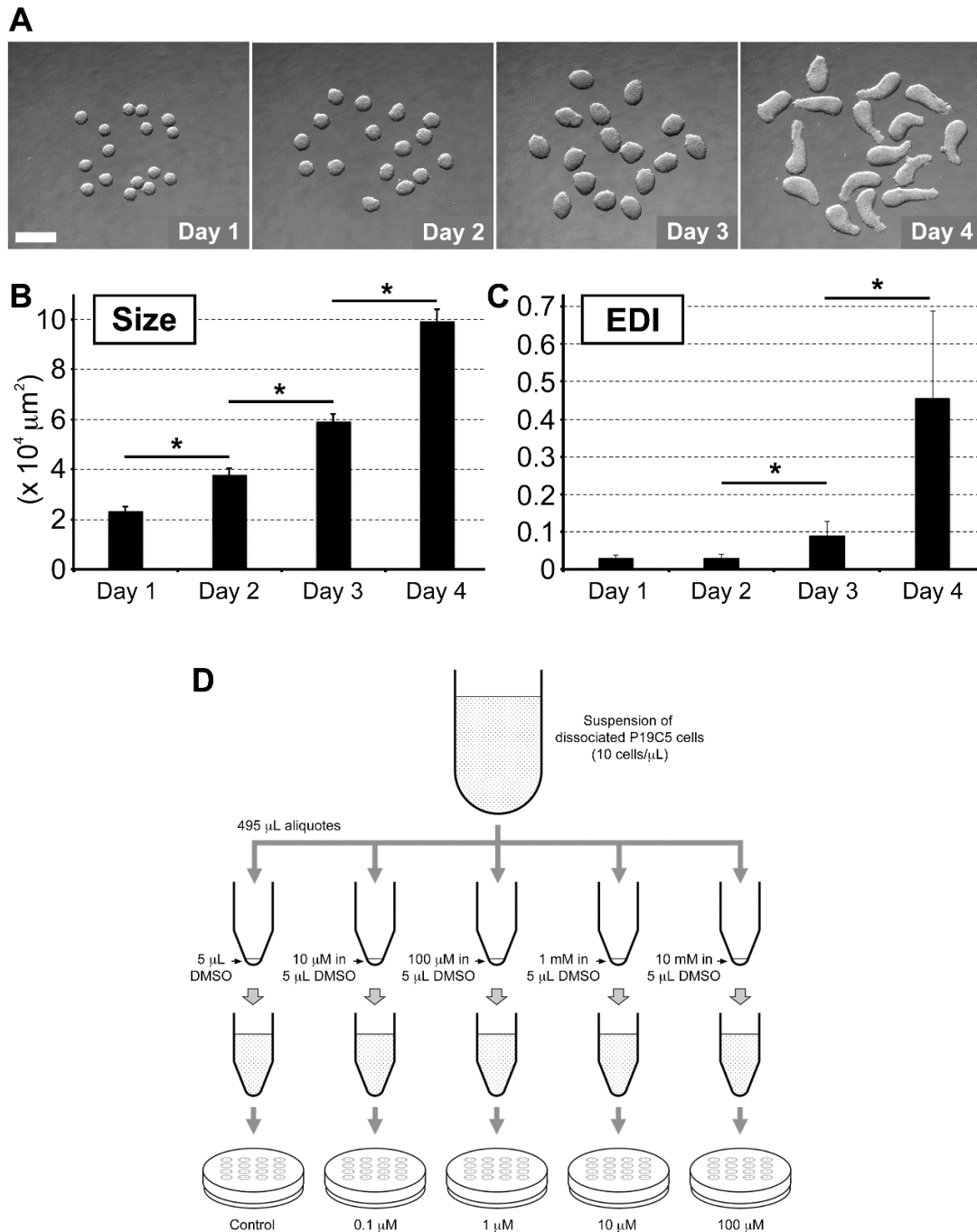
Developmental toxicants that exhibit species-specific effects may also be misclassified by the P19C5 system, which is composed of mouse stem cells. Although thalidomide is a well-known teratogen in humans, the limb malformations characteristic of thalidomide-induced developmental toxicity are essentially absent in rodents (Brent, 1964). Thalidomide binds and inhibits the cereblon protein, which appears to be the primary cause of its teratogenic effects (Ito *et al.*, 2010). However, homozygous cereblon-knockout mice exhibit only behavioral problems with no limb defects (Rajadhyaksha *et al.*, 2012). Thus, failure of thalidomide to impact P19C5 morphogenesis may be reflective of species-specific teratogenicity. To detect developmental toxicants that specifically affect human embryos, *in vitro* morphogenesis models composed of human cells may be necessary. Many researchers have adopted human ES cells (hESC) as more human-relevant *in vitro* models to assess developmental toxicity of chemical agents, and some hESC systems have been shown to detect the effect of thalidomide (Kameoka *et al.*, 2014; Xing *et al.*, 2015). Previously, we established a culture condition to induce elongation morphogenesis in EBs made of hESCs (Lau and Marikawa, 2014). It is of particular interest to examine whether thalidomide affects morphogenesis of these hESC EBs in future studies.

Although considered safe to be used during pregnancy, both doxylamine and nystatin impaired P19C5 EB development. This apparent misclassification highlights another limitation of the experimental scheme that was adopted in the present study, i.e., methodical testing of a set of predetermined concentrations. The concentration range of 0.1 to 100  $\mu$ M may not reflect realistic levels of exposure of certain drugs to the developing human embryo. As an antifungal agent, nystatin is administered topically or orally and is minimally absorbed into circulation. Thus, the chance of embryonic exposure to high levels of nystatin is negligible, which likely contributed to the low pregnancy risk classification of this drug. However, nystatin is also known to be highly toxic when administered intravenously. The adverse impact on EB development may have been reflecting the toxic potential of nystatin at high plasma concentrations, which are unlikely to occur in pregnant women who use nystatin only topically or orally.

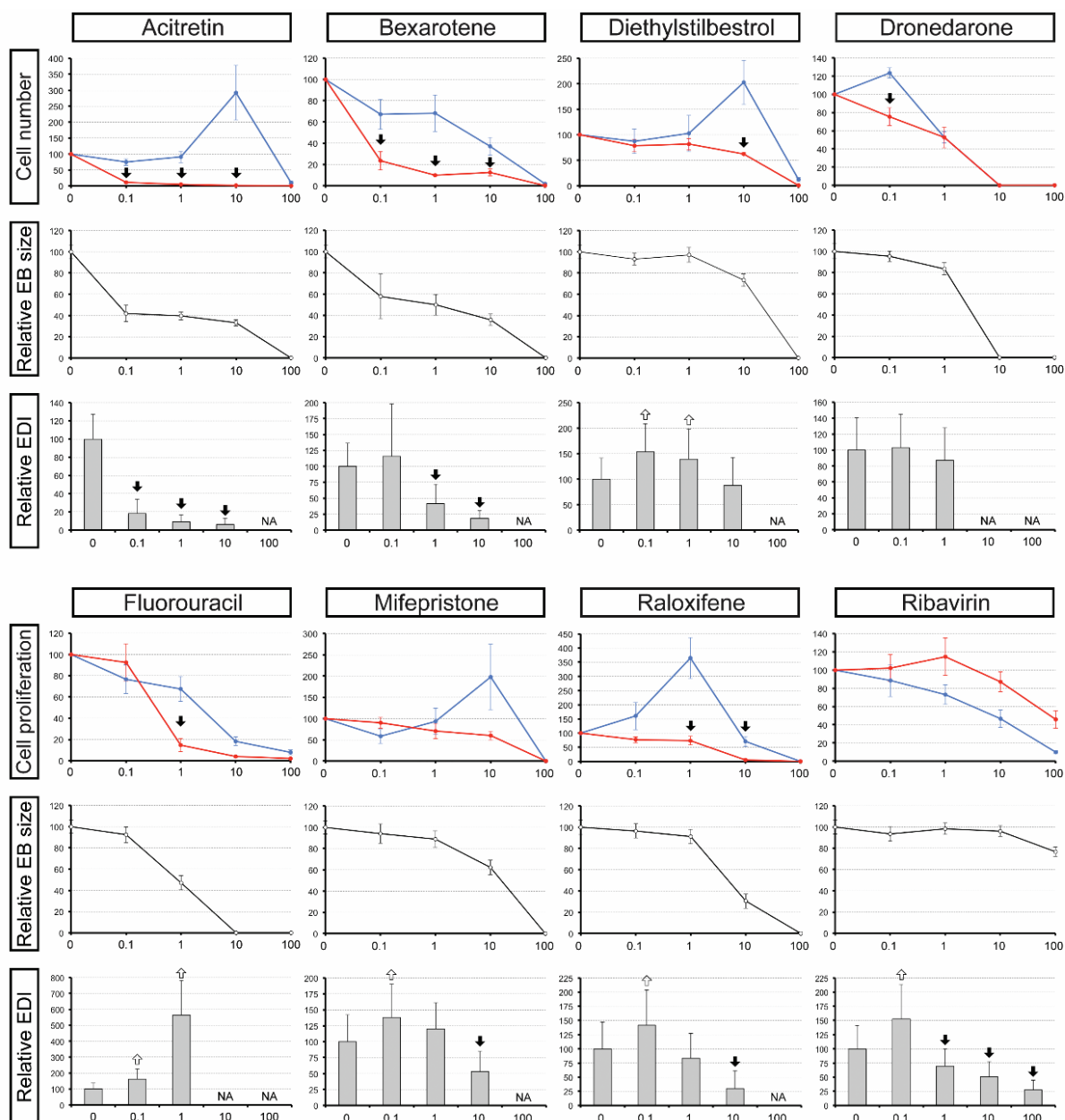
Also, the concentration of doxylamine that affected P19C5 cell number and EDI was 100  $\mu$ M, which is far higher than the plasma concentration found after routine doses in patients (Videla *et al.*, 2013). Therefore, in future studies to validate the P19C5 EB system, it is important to consider realistic exposure levels for each drug — a concept that is extensively discussed in recent studies by Daston *et al.* (2010, 2014) as a critical feature of effective developmental toxicant assays.

As a potential tool for developmental toxicity testing, the P19C5 EB system has several practical advantages over other *in vitro* methods. First, the P19C5 system requires only four days from the start of hanging drop culture to morphometric analysis. In contrast, the standard ESTc protocol requires a total of 10 days of culture to observe beating of cardiomyocytes (Spielmann *et al.*, 1997; Seiler *et al.*, 2006). Modifications of ESTc have shortened the duration of the assay by analyzing specific cardiac marker gene expression, but still require six days of culture (Suzuki *et al.*, 2011). Secondly, the procedure for P19C5 EB culture is simple, as it takes only one series of hanging drop culture without any change of culture medium or culture format. This is an advantage over other EST protocols that require changes in culture medium and/or format in order to promote differentiation of specific cell types. Thirdly, simple tracing of EB circumferences in the ImageJ program can provide data on multiple morphometric parameters, including the size and EDI, which are altered by developmental toxicity of test compounds. Each morphometric parameter is likely to be affected by specific aspects of EB development. For example, changes in cell proliferation would impact EB size whereas alteration in cell migration and adhesion would affect EDI. Thus, morphometric analyses can provide mechanistic insight into the action of developmental toxicants without intensive labor or techniques. Lastly, the P19C5 EB system may be more economical. Culturing the P19C5 cell line is less expensive than culturing ES cell lines because it does not require supplementation of Leukemia Inhibitory Factor for the maintenance of undifferentiated state. Furthermore, drug impact in the P19C5 system can be assessed visually using the ImageJ program, which is freely available from the National Institutes of Health. These practical advantages do not presume effectiveness or applicability of the P19C5 EB systems. However, the establishment of faster, simpler, and cheaper methods is necessary for pharmaceutical companies to conduct thorough and extensive developmental toxicity tests on the hundreds or thousands of new therapeutic compounds.

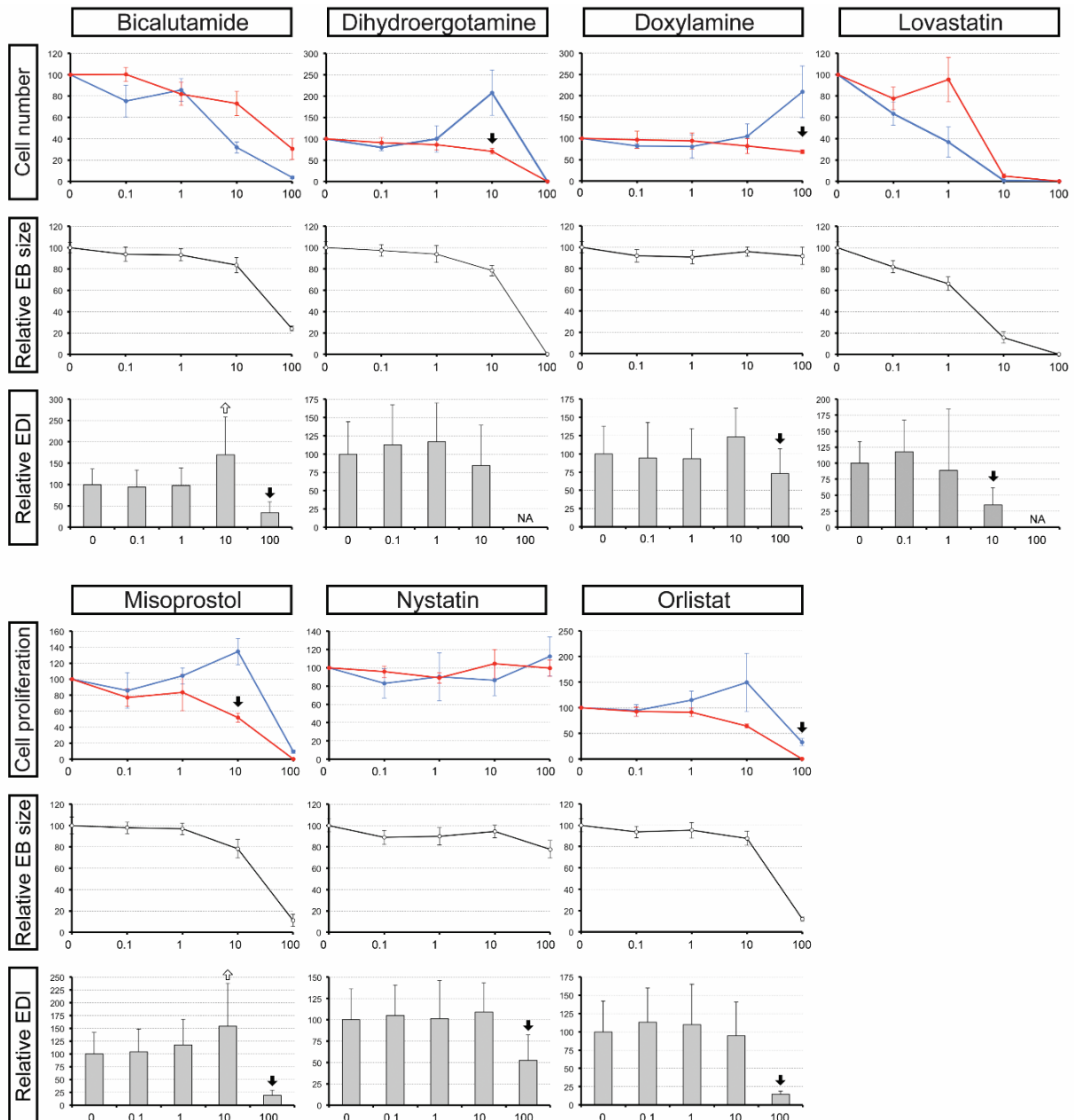
Though further validation studies are needed, the present study shows that the P19C5 EB system is a promising non-animal test for developmental toxicants, because it has potential of high accuracy, offers economic practicality, and is based on realistic embryologic processes of growth and morphogenesis that are often targeted in various birth defects.



**Figure 2.1.** *In vitro* morphogenesis of P19C5 embryoid bodies (EBs). (A) Photographs of control EBs over four days of culture to demonstrate the time course of morphological transformation. Scale bar = 500  $\mu\text{m}$ . (B, C) Temporal changes in the EB size (B) and EDI (C). Graphs are mean + standard deviation ( $n = 44$ ). Asterisks indicate significant differences between two adjacent groups ( $p < 0.01$ ; two-sample t-test). (D) The experimental scheme to examine morphogenetic impact of the therapeutic drugs.

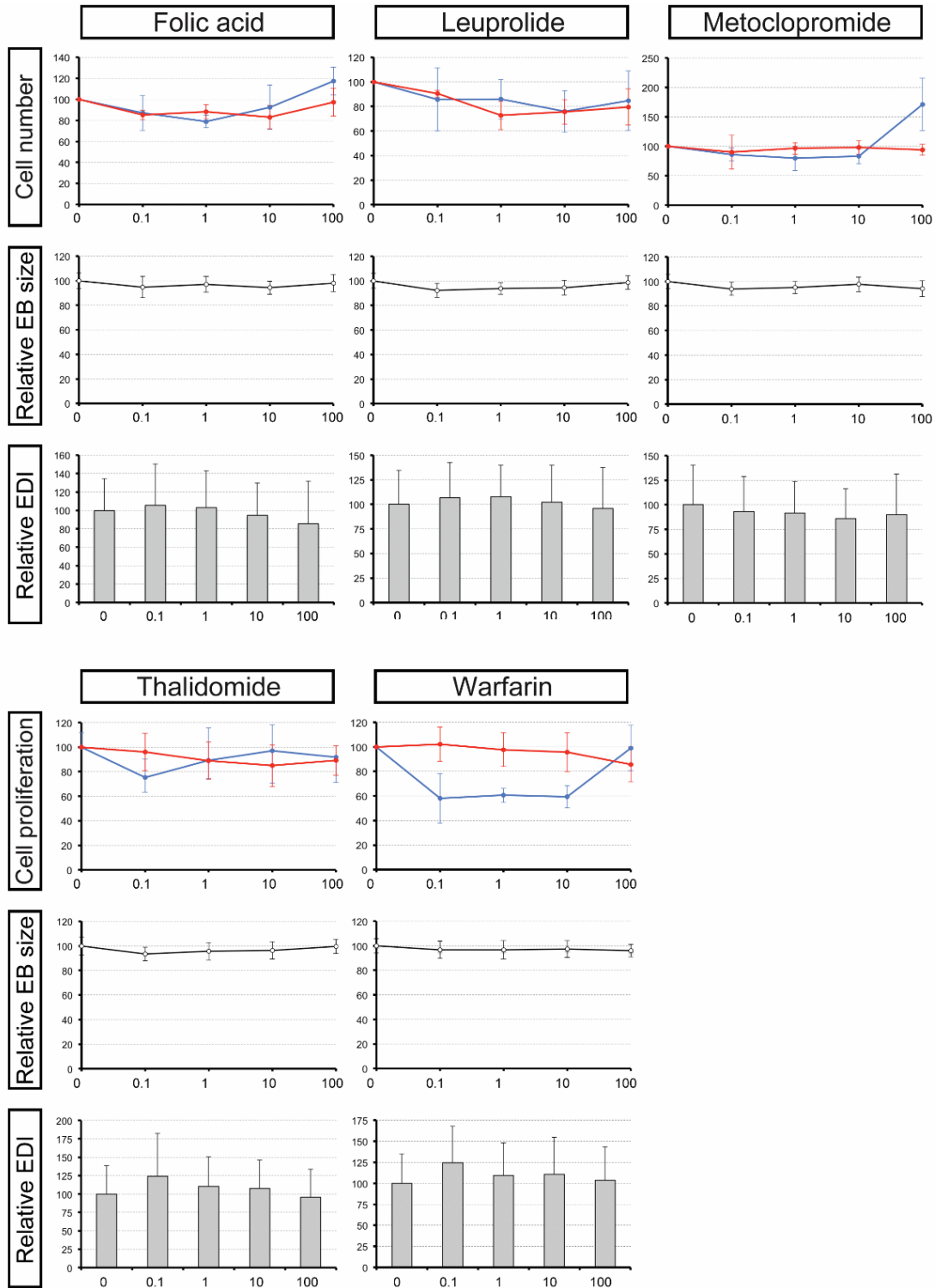


**Figure 2.2.** Drugs that impacted P19C5 cells with high potency. Three types of graphs are shown for each drug. Relative values are calculated as a percentage of the corresponding set's control value. All control values are shown as 100. Top: Relative cell numbers of P19C5 (red line) and NIH/3T3 (blue line) after 4 days of mono-layer culture in the presence of drug. Mean  $\pm$  standard deviation is presented for each cell line ( $n = 4$ ). Downward black arrows indicate significant reductions in P19C5 cell number relative to NIH/3T3 cell number ( $p < 0.01$ ). Middle: Relative size of P19C5 EBs after 4 days of hanging drop culture in the presence of drug. Mean  $\pm$  standard deviation is presented as the data points and error bars respectively ( $n = 45$  to  $48$ ). Values of "0" for relative EB size indicate death of EBs. Bottom: Relative EDI (Elongation Distortion Index) of P19C5 EBs after 4 days of hanging drop culture. Mean  $\pm$  standard deviation is presented as each column's height and error bar respectively ( $n = 45$  to  $48$ ). Downward black arrows indicate significant reductions in relative EDI whereas upward white arrows indicate significant increase in relative EDI ( $p < 0.01$ ). NA: Not applicable due to death of EBs. Horizontal axes represent drug concentrations in micromolar.



**Figure 2.3.** Drugs that impacted P19C5 cells with mild potency. Relative values are calculated as a percentage of the corresponding set's control value, and all control values are shown as 100. Top: Relative cell numbers of P19C5 (red line) and NIH/3T3 (blue line) after 4 days of mono-layer culture in the presence of drug. Mean  $\pm$  standard deviation is presented for each cell line ( $n = 4$ ). Downward black arrows indicate significant reductions in P19C5 cell number relative to NIH/3T3 cell number ( $p < 0.01$ ). Middle: Relative size of P19C5 EBs after 4 days of hanging drop culture in the presence of drug. Mean  $\pm$  standard deviation is presented as the data points and error bars respectively ( $n = 45$  to  $48$ ). Bottom: Relative EDI (Elongation Distortion Index) of P19C5 EBs after 4 days of hanging drop culture. Mean  $\pm$  standard deviation is presented as each column's height and error bar respectively ( $n = 45$  to  $48$ ). Downward black arrows indicate significant reductions in relative EDI whereas upward white arrows indicate significant increase in relative EDI ( $p < 0.01$ ). NA: Not applicable due to death of EBs. Horizontal axes represent drug concentrations in micromolar.





**Figure 2.4.** Drugs that did not impact P19C5 cells. Top: Relative cell numbers of P19C5 (red line) and NIH/3T3 (blue line) after 4 days of mono-layer culture in the presence of drug. Mean  $\pm$  standard deviation is presented for each cell line ( $n = 4$ ). Middle: Relative size of P19C5 EBs after 4 days of hanging drop culture in the presence of drug. Mean  $\pm$  standard deviation is presented ( $n = 45$  to  $48$ ). Bottom: Relative EDI (Elongation Distortion Index) of P19C5 EBs after 4 days of hanging drop culture. Mean  $\pm$  standard deviation is presented ( $n = 45$  to  $48$ ). Horizontal axes represent drug concentrations in micromolar.

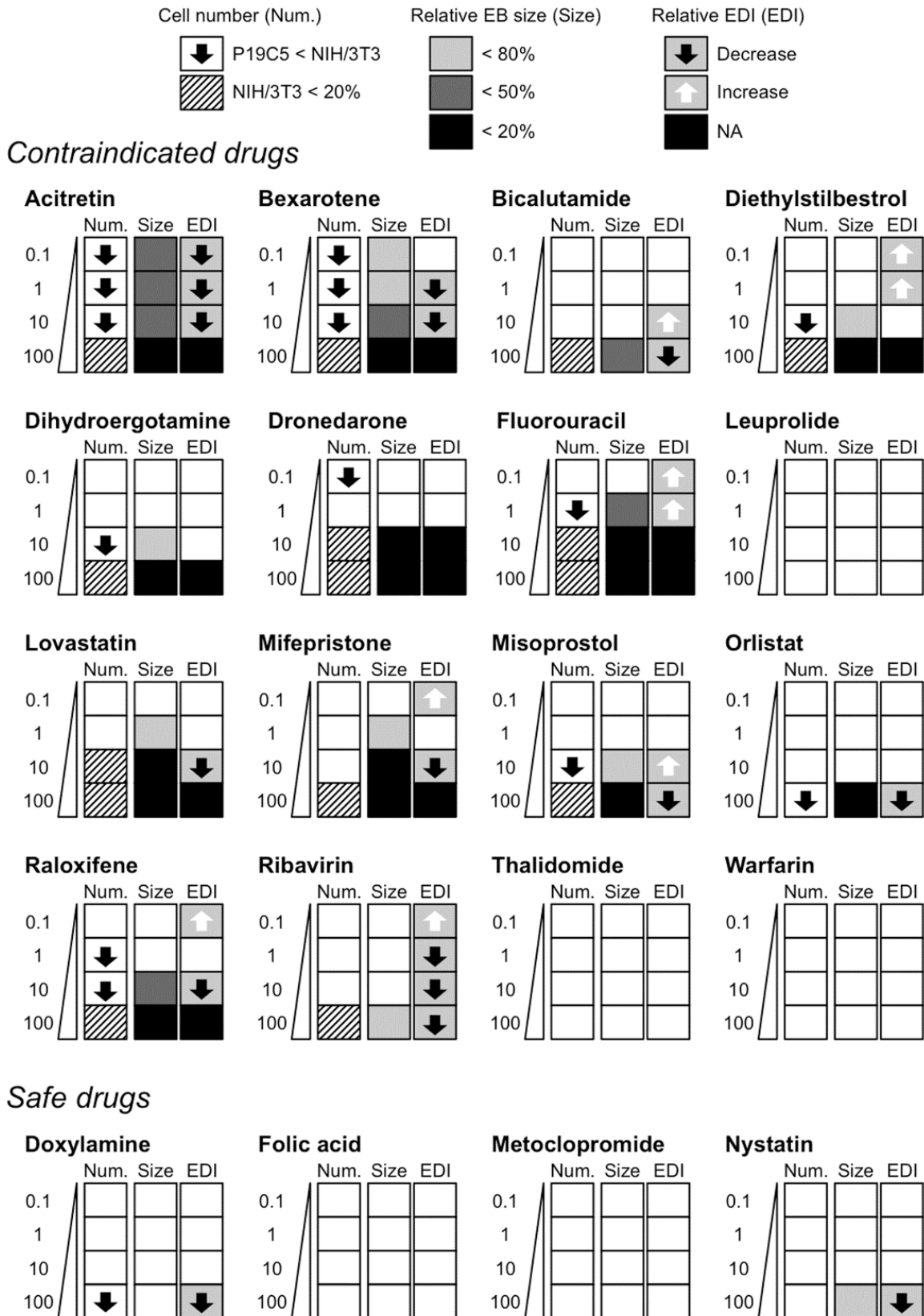
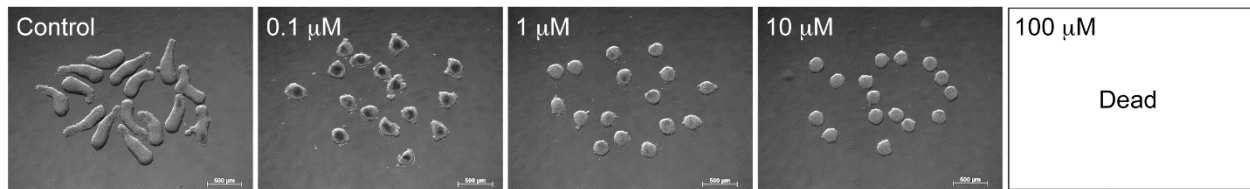
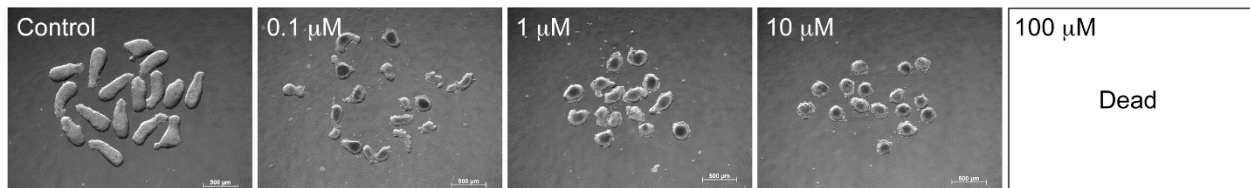


Figure 2.5. Summary diagram of drug impacts. See text for details.

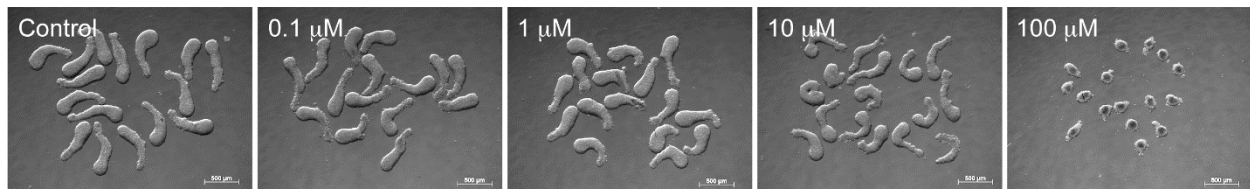
## A Acitretin



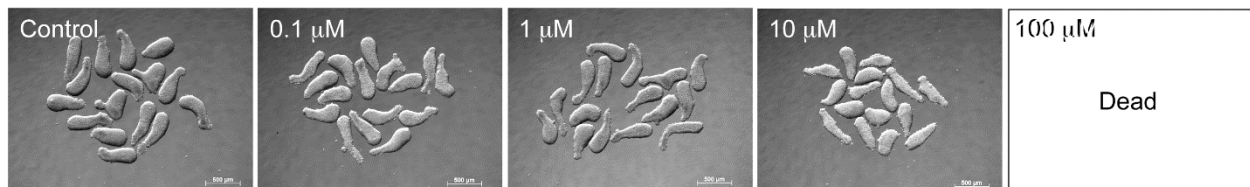
## B Bexarotene



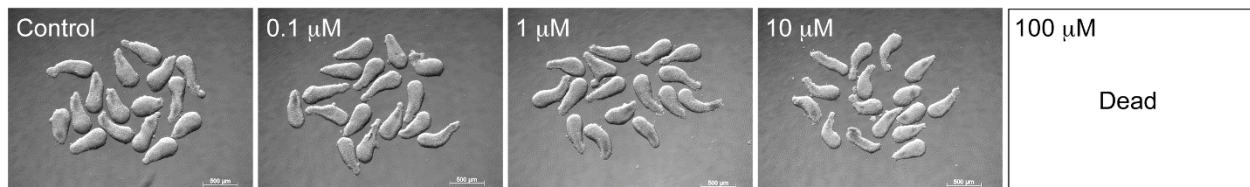
## C Bicalutamide



## D Diethylstilbestrol

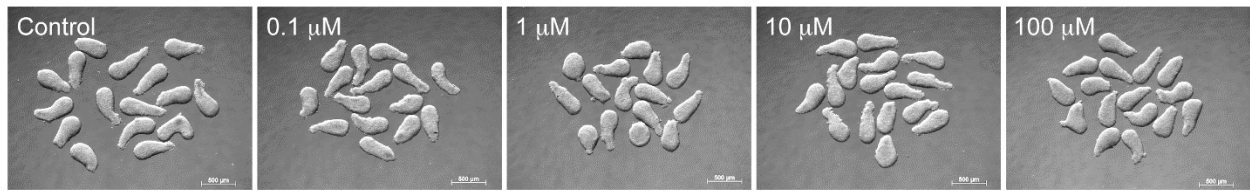


## E Dihydroergotamine

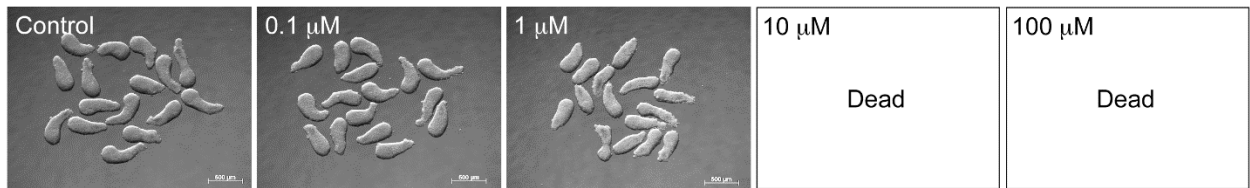


**Figure 2.6.** Example images of P19C5 EBs after four days of treatment in hanging drop culture.

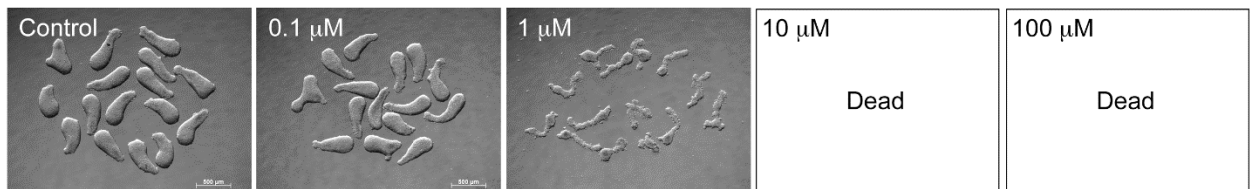
## F Doxylamine



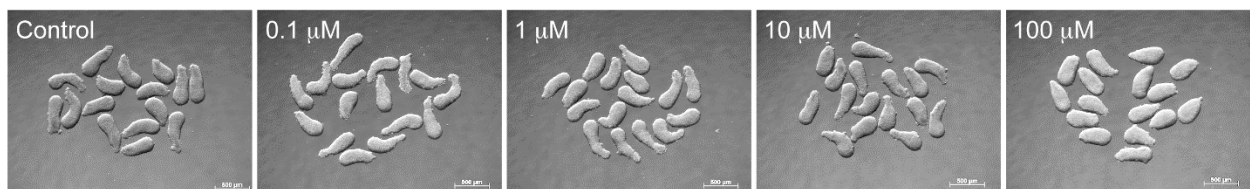
## G Dronedarone



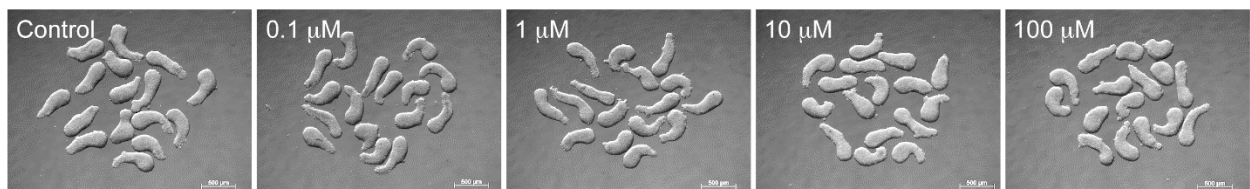
## H Fluorouracil



## I Folic acid



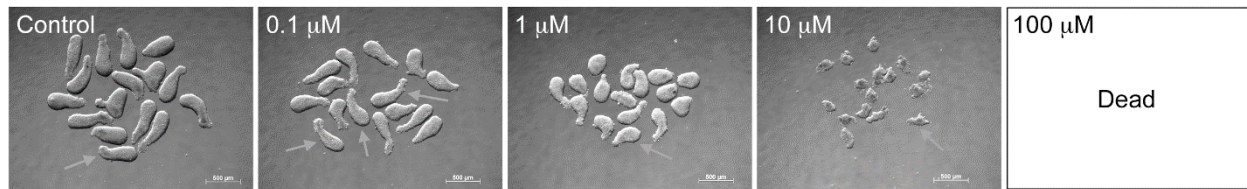
## J Leuprolide



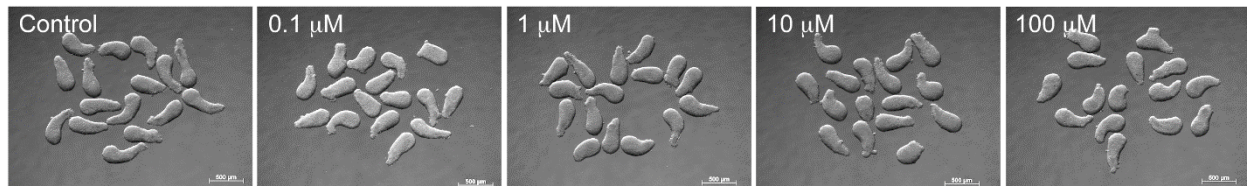
**Figure 2.6. (Continued).** Example images of P19C5 EBs after four days of treatment in hanging drop culture.



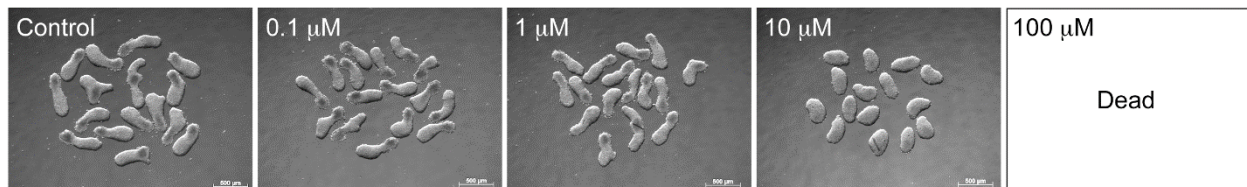
## K Lovastatin



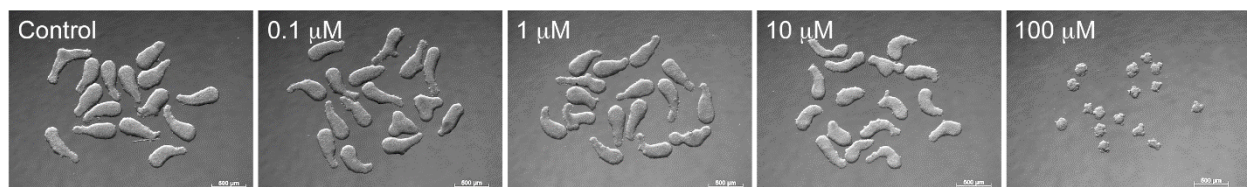
## L Metoclopramide



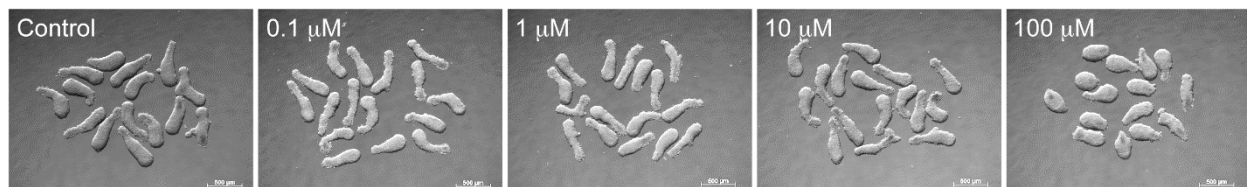
## M Mifepristone



## N Misoprostol

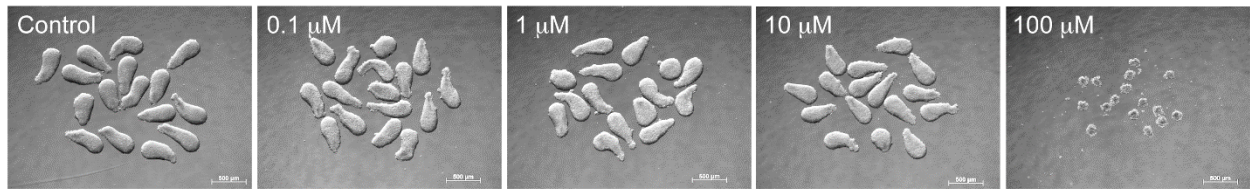


## O Nystatin

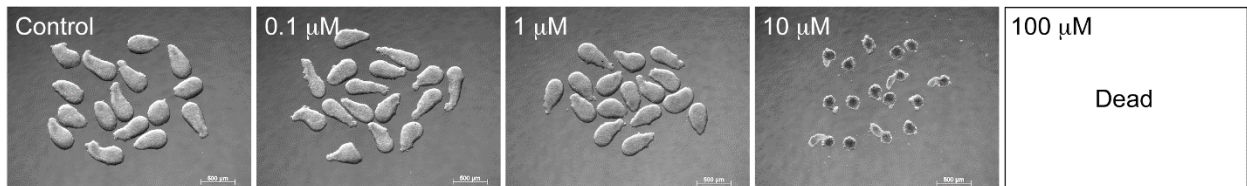


**Figure 2.6. (Continued).** Example images of P19C5 EBs after four days of treatment in hanging drop culture.

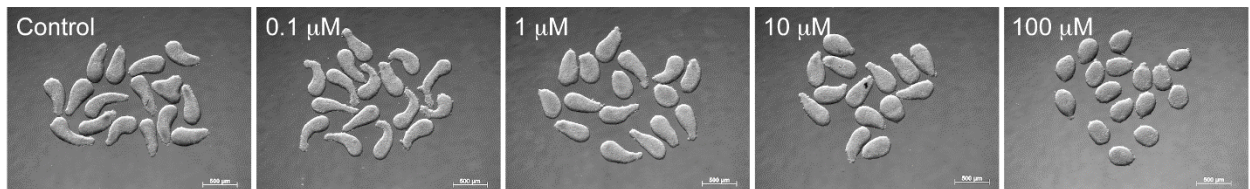
## P Orlistat



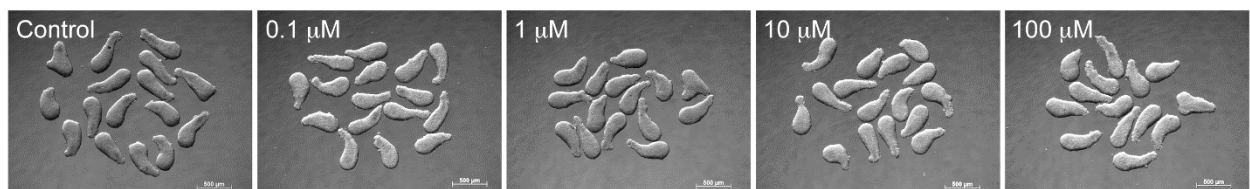
## Q Raloxifene



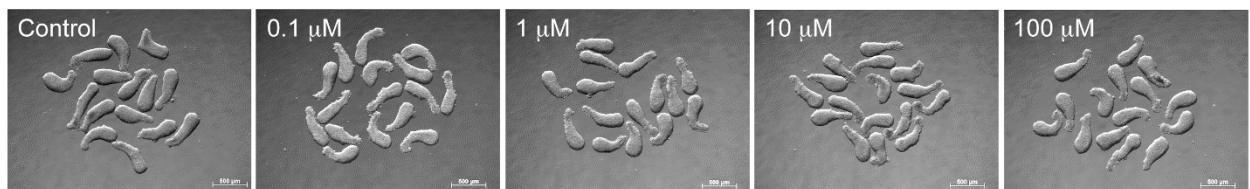
## R Ribavirin



## S Thalidomide



## T Warfarin



**Figure 2.6. (Continued).** Example images of P19C5 EBs after four days of treatment in hanging drop culture.

**Table 2.1.** Compounds used in the Category X validation study.

<b>Drug Name (CAS RN<sup>*1</sup>)</b>	<b>Pregnancy Risk</b>	<b>Therapeutic Target</b>	<b>Chemical Properties</b>	<b>References<sup>*2</sup></b>
Acitretin (CAS 55079-83-9)	Contraindicated	Dermatologic diseases	Synthetic retinoid (RAR receptor agonist)	Barbero <i>et al.</i> , 2004; Geiger <i>et al.</i> , 1994
Bexarotene (CAS 153559-49-0)	Contraindicated	Cutaneous T-cell lymphoma	Synthetic retinoid (RXR receptor agonist)	Lu <i>et al.</i> , 1997; Elmazar and Nau, 2004
Bicalutamide (CAS 90357-06-5)	Contraindicated	Prostate cancer, Hirsutism	Non-steroidal anti-androgen	Sadar <i>et al.</i> , 1999; Cockshott, 2004
Diethylstilbestrol (CAS 56-53-1)	Contraindicated	Miscarriage; Premature deliveries	Non-steroidal estrogen	Mahawong <i>et al.</i> , 2014; Reed and Fenton, 2013
Dihydroergotamine (CAS 511-12-6)	Contraindicated	Migraine	Synthetic ergot alkaloid	Bérard and Kori, 2012; Hohmann and Künzel, 1992
Doxylamine (CAS 469-21-6)	Safe	Allergy	First generation antihistamine	Gilboa <i>et al.</i> , 2014; Slaughter <i>et al.</i> , 2014
Dronedarone (CAS 141626-36-0)	Contraindicated	Arrhythmia	Multi-channel blocker	Marzocchi and Lombardi, 2011; Eskes and Wiersinga, 2009
Fluorouracil (CAS 51-21-8)	Contraindicated	Cancer	Thymidylate synthase inhibitor	Kuwagata <i>et al.</i> , 1998; Murthy <i>et al.</i> , 2014
Folic Acid (CASRN 59-30-3)	Safe	Used as dietary supplement	Vitamin B9	van Mil <i>et al.</i> , 2010; Charles <i>et al.</i> , 2005
Leuprolide (CAS 53714-56-0)	Contraindicated	Endometriosis; Prostate cancer	LHRH receptor agonist	Abu-Heija <i>et al.</i> , 1995; Cahill <i>et al.</i> , 1994
Lovastatin (CAS 75330-75-5)	Contraindicated	Hyper-cholesterolemia	HMG-CoA reductase inhibitor	Lankas <i>et al.</i> , 2004; Godfrey, 2012
Metoclopramide (CAS 364-62-5)	Safe	Nausea	Dopamine receptor antagonist	Pasternak <i>et al.</i> , 2013; Matok and Perlman, 2014
Mifepristone (CAS 84371-65-3)	Contraindicated	Used as abortifacient	Progesterone receptor antagonist	Chen <i>et al.</i> , 2014; Blanch <i>et al.</i> , 1998
Misoprostol (CAS 59122-46-2)	Contraindicated	Gastric ulcers	Prostaglandin E1 analog	Bos-Thompson <i>et al.</i> , 2008; Schüller <i>et al.</i> , 1999
Nystatin (CAS 1400-61-9)	Safe	Fungal infection	Ergosterol-binding	King <i>et al.</i> , 1998; Larson <i>et al.</i> , 2000
Orlistat (CAS 96829-58-20)	Contraindicated	Obesity	Lipase inhibitor	Källén, 2014; Browne <i>et al.</i> , 2006
Raloxifene (CAS 84449-90-1)	Contraindicated	Osteoporosis; Breast cancer	Selective estrogen receptor modulator	Gizzo <i>et al.</i> , 2013; Komm and Mirkin, 2014
Ribavirin (CAS 36791-04-5)	Contraindicated	Viral infection	Purine analog	Roberts <i>et al.</i> , 2010; De Santis <i>et al.</i> , 2003
Thalidomide (CAS 50-35-1)	Contraindicated	Leprosy; Multiple myeloma	Immuno-modulatory agent	Vargesson, 2009; Ito and Handa, 2012
Warfarin (CAS 81-81-2)	Contraindicated	Used as blood thinner	Vitamin K epoxide reductase inhibitor	Starling <i>et al.</i> , 2012; Mehndiratta <i>et al.</i> , 2010

<sup>\*1</sup> CAS RN: Chemical Abstracts Service Registration Number

<sup>\*2</sup> References that are relevant to pregnancy risk and/or mechanisms of actions

## CHAPTER 3. EXPOSURE-BASED VALIDATION OF AN *IN VITRO* GASTRULATION MODEL FOR DEVELOPMENTAL TOXICITY ASSAYS

### 3.1 - Introduction

*In utero* exposure to developmental toxicants can interfere with the normal course of embryonic development and result in abnormalities, such as death, growth retardation and malformations. Many medications, herbicides and industrial byproducts are known to be developmental toxicants (Schardein and Macina, 2006). However, many more compounds have not been sufficiently evaluated for potential developmental toxicity, so that continued efforts are required to identify adverse impact on embryos and to minimize harmful exposures to women of child-bearing age. *In vitro* model systems, particularly those using pluripotent stem cells, are promising screening tools for developmental toxicants. Pluripotent stem cells, such as embryonic stem (ES) cells, can differentiate into multiple tissue types *in vitro* and display embryo-like properties. Thus, pluripotent stem cells can recapitulate certain aspects of embryonic development and serve as *in vitro* models to demonstrate the impacts of developmental toxicants. ES cell tests, or ESTs, evaluate the developmental toxicity of compounds based on their inhibitory effects on ES cell differentiation (Riebeling *et al.*, 2012; Theunissen *et al.*, 2012), but each individual EST system recapitulates only a limited aspect of embryonic development, i.e., differentiation of cardiomyocytes, neurons, or osteoblasts. Thus, it is more likely that a panel of complementary *in vitro* systems — each representing distinct aspects of embryonic development — can comprehensively screen a broad range of developmental toxicants.

To assemble a panel of complementary assays, each assay must undergo a series of validation studies using known developmental toxicants to assess its applicability and limitations. However, selection of proper reference compounds to validate *in vitro* screening assays has been challenging, partly because developmental impact of compounds varies depending on timing, dose and duration of exposure to embryos (Friedman, 2010; Jelínek, 2005). For example, a given compound can exert developmental toxicity when exposed at a high dose, but the same compound may pose no risk at lower doses (Daston *et al.*, 2010). Therefore, dichotomic designation of compounds as developmental toxicants or non-toxicants is ineffective in properly validating assays. To that end, Daston and colleagues have proposed an “exposure-based validation list” for developmental toxicity screening assays (Daston *et al.*, 2014). This list (referred hereafter as the Daston list) consists of 20 “positive exposures” known to cause embryo-fetal death or structural malformation in rats, and 19 “negative exposures”, which have no adverse impacts on rat embryo development. Eleven of these compounds demonstrate both a positive exposure at a high concentration and a negative exposure at a low concentration. The Daston list is a significant step forward in validating *in vitro* assays for developmental toxicity screening.



The objective of the present study is to validate the *in vitro* gastrulation model of P19C5 stem cells using the compounds on the Daston list. P19C5 cells are mouse embryonal carcinoma stem cells that possess properties similar to the pluripotent epiblast lineage of post-implantation embryos. Three-dimensional culture of P19C5 cell aggregates results in spontaneous differentiation of mesendoderm along with steady increase in size and axial elongation (Lau and Marikawa, 2014). This morphogenetic transformation of P19C5 cell aggregates, or embryoid bodies (EBs), resembles gastrulation, the morphogenetic process of early embryonic development that creates the germ layers and elongated body shape along the anterior-posterior axis. Growth and axial elongation of the P19C5 gastrulation model are impaired by pharmacological inhibitors of the major developmental signals that are crucial for embryo body patterning, namely Wnt, Nodal, Fgf and retinoic acid signaling pathways (Li and Marikawa, 2015). The morphogenesis of P19C5 EBs is also sensitive to various therapeutic drugs under the Food and Drug Administration (FDA) Pregnancy Risk Category X, i.e., those contraindicated for use during pregnancy (Warkus *et al.*, 2016). Furthermore, valproic acid (VPA) at 0.8 mM, one of the positive exposures in the Daston list, affects the P19C5 morphogenesis, such that VPA-treated EBs are smaller, distorted, and less elongated (Li and Marikawa, 2016). These studies suggest that the P19C5 gastrulation model can serve as an effective *in vitro* tool to detect developmental toxicants.

Here, we validated the P19C5 gastrulation model using the 34 exposures compiled in the Daston list. While the morphology-based assay correctly classified many of the Daston exposures, the present study also revealed some limitations of the assay based on the cases where exposures were misclassified.

## 3.2 - Materials and methods

### 3.2.1 TEST COMPOUNDS

Compounds of the Daston list (Daston *et al.*, 2014) used in the present study are shown in Table 1. Four compounds in the Daston list, namely HEPP (positive exposure), SB-209770 (positive and negative exposures), tapentadol (negative exposure), and valproic acid (positive exposure), were not evaluated for the following reasons. HEPP and SB-209770 were unavailable from major chemical suppliers. Tapentadol was available from one supplier (Sigma-Aldrich, St Louis, Missouri) as a solution of 1.0 mg/mL dissolved in methanol (Catalog Number T-058). However, to achieve the concentration indicated in the Daston list (252 mg/L), about 25% of the culture medium would be methanol, which is not compatible with our assay (see below). The morphogenetic effect of valproic acid on the P19C5 gastrulation model has been evaluated previously, including the concentration cited in the Daston list (0.8 mM) (Li and Marikawa, 2016). Note that the concentration of desloratadine indicated in the Daston study (1.5 mM) (Daston *et al.*, 2014) is apparently a typographical error, as the original study referenced therein (FDA, 2001) indicates 1.5  $\mu$ M instead. Accordingly, desloratadine was evaluated at 1.5  $\mu$ M as a negative exposure in the present study.

### 3.2.2 CELL CULTURE

P19C5 cells, a subline of P19 mouse embryonal carcinoma cell line (Lau and Marikawa, 2014), were propagated in culture medium (Minimum Essential Medium Alpha with nucleosides and GlutaMAX Supplement [LifeTechnologies, Carlsbad, California], 2.5% fetal bovine serum, 7.5% newborn calf serum, 50 units/mL penicillin, and 50  $\mu$ g/mL streptomycin). Embryoid bodies (EBs) of P19C5 cells were generated according to the method previously described for P19 cell aggregates (Marikawa *et al.*, 2009). Briefly, P19C5 cells were fully dissociated with Trypsin-EDTA, and suspended in culture medium containing 1% dimethyl sulfoxide (DMSO) at the density of 10 cells/ $\mu$ L with or without specific amount of a test compound (Fig. 3.1). Drops (20  $\mu$ L each) of cell suspension were spotted on the inner surface of Petri dish lids for hanging drop culture. NIH/3T3 (derived from mouse embryonic fibroblast), HEK293 (derived from human embryonic kidney), and JEG3 (derived from human choriocarcinoma) were obtained from the American Type Culture Collection (Manassas, Virginia), and propagated in culture medium (Dulbecco's Modified Eagle Medium, GlutaMAX Supplement [LifeTechnologies], 10% fetal bovine serum, 50 units/mL penicillin, and 50  $\mu$ g/mL streptomycin).

### 3.2.3 IMAGE ANALYSIS

Embryoid bodies (EBs) were removed from hanging drops and grouped together for photography using an AxioCam MRm digital camera connected to an Axiovert 200 microscope with Hoffman

modulation-contrast optics (Carl Zeiss, Thornwood, NY). Image files were converted to JPEG format and opened in ImageJ (<http://rsb.info.nih.gov/ij>). Morphological parameters of individual EBs were measured on ImageJ by tracing their circumference. In the previous studies (Lau and Marikawa, 2014; Li and Marikawa, 2015; Warkus *et al.*, 2016), tracing of EB circumference was performed manually using the polygon selection tool. In the present study, however, we formulated a series of program operations to enable tracing of EB circumference in a faster and less laborious fashion. The details of the program operations are described in Appendices A and B. Briefly, the first set of operations (e.g., Image Calculator, Find Edges, Brightness/Contrast and Binary) was applied to convert the inside of EBs into solid black while the background into white. The second set of operations (e.g., Wand Tool and Fit Splines) was applied to detect the outlines of blackened areas, corresponding to the circumference of EBs. The third set of operations (e.g., Measure) was applied to selected regions of interest (ROI) to determine their morphological parameters, namely area and circularity ( $= 4 \times \pi \times \text{area} / \text{perimeter}^2$ ). Measurements were exported to Microsoft Excel, where Elongation Distortion Index (EDI =  $1 / \text{circularity} - 1$ ) was calculated. As described previously, area was used as a proxy for the size of EB, whereas EDI was used to gauge the extent of EB axial elongation (Warkus *et al.*, 2016).

### **3.2.4 VIABILITY ASSAY**

The impact of drugs on cell proliferation and viability was evaluated using CellTiter-Glo Luminescent Cell Viability Assay system (Promega, Madison, Wisconsin), which determines the number of live cells in culture by measuring the amount of ATP as a quantitative proxy for the number of metabolically active cells. P19C5, NIH/3T3, HEK293 and JEG3 cells were seeded in 96-well plates at the density of 100, 250, 250 and 500 cells/well, respectively, and were cultured in the corresponding medium (100  $\mu\text{l}$ /well) supplemented with 1% DMSO with or without specific amount of a test compound. After 4 days of culture, cells were treated with CellTiter-Glo Reagent for measurement of luminescence, as a readout of ATP amount, according to the manufacturer's instruction (Promega), using Gene Light 55 Luminometer (Microtech, Chiba, Japan). Cell seeding density was optimized through a series of pilot experiments, to confirm that cell numbers at the end of four days of culture were proportionate to intensities of luminescence. Relative cell number was calculated based on ratio of the luminescence intensity in compound-treated cells to that in non-treated cells of the same set of experiments. For each compound exposure, three sets of experiments were conducted as biological replicates.

### **3.2.5 STATISTICAL ANALYSES**

Experiments to assess morphological impacts of compound exposures on EB morphology were conducted in three biological replicates using different collections of cell suspensions. For each replicate, 16 hanging drops were generated per exposure (a specific concentration of compound) in parallel with 16

control (i.e., no compound) hanging drops (Fig. 3.1). Area and EDI of individual Day 4 EBs were normalized against the average of control EBs, and defined as relative area and relative EDI, respectively, expressed in percentages (i.e., averages of relative area and relative EDI of control EBs are 100%). Data from 3 replicates were compiled, and their averages are shown with 95% confidence intervals (Fig. 3.2). Thus, a total of 46 to 48 EBs were scored for each exposure (1 or 2 EBs were occasionally lost or damaged during operations). To verify that observed effects on EBs were statistically significant, two-sample t-test was performed between compound-treated group and the matching control group. All morphological impacts that were defined as adverse in the present study were statistically significant ( $P < 0.01$ ).

## 3.3 - Results

### 3.3.1 EXPERIMENTAL DESIGN

The overall experimental scheme is shown in Fig. 3.1. In the present study, we examined several concentrations for each compound in addition to those in the Daston list (Daston *et al.*, 2014). Particularly, when Daston's positive exposures had no apparent effect on P19C5 EBs, higher concentrations were also evaluated. EBs were observed daily for survival and integrity. On the fourth day of culture (Day 4), EBs were photographed and analyzed using the ImageJ program to measure the morphometric parameters, namely area and EDI (see Materials and Methods). Impact of each compound exposure was assessed based on changes in area or EDI relative to control EBs. In the present study, a compound exposure was classified as having an "adverse" effect on EB morphogenesis when it caused either of the following three outcomes: (1) degeneration (i.e., death) of EBs at any time point in culture or cell aggregation failure, (2) a reduction in the average area by more than 20% relative to control EBs, and (3) a decrease or increase in the average EDI by more than 40% relative to control EBs. Note that an increase in the average EDI accompanied by an increase in the average area was not classified as "adverse", because such condition suggests "promotion" of EB growth and morphogenesis rather than "impairment" (see Discussion).

### 3.3.2 EFFECTS OF THE POSITIVE AND NEGATIVE EXPOSURES ON EB MORPHOGENESIS

Morphogenetic impacts of compound exposures on P19C5 EBs are shown in Fig. 3.2 and summarized in Tables 2 and 3 with respect to specific concentrations indicated in the Daston list. Sample images of EBs, namely those that were distinctly affected by exposures, are shown in Fig. 3.3. Based on the classification criteria described above, 58.8% (10 out of 17) of the positive exposures tested had adverse impact on P19C5 EBs (Table 2.2). Specifically, acetazolamide (120  $\mu$ M) reduced relative area by 25% and reduced relative EDI by 45%; all-*trans* retinoic acid (ATRA; 200 nM) reduced relative area by 60% and reduced relative EDI by 90%; caffeine (325  $\mu$ M) increased relative EDI by 70%; hydroxyurea (350  $\mu$ M) caused death; mono(2-ethylhexyl) phthalate (MEHP; 150  $\mu$ M) reduced relative area by 25%; methoxyacetate (5 mM) reduced relative area by 25% and reduced relative EDI by 90%; methylmercury (5  $\mu$ M) caused death; nilotinib (28  $\mu$ M) prevented cell aggregation (without apparent cell death); ramelteon (80  $\mu$ M) increased relative EDI by 45%; and salicylic acid (3 mM) reduced relative area by 60%. In contrast, only 17.6% (3 out of 17) of the negative exposures exhibited adverse effects on EB development (Table 3.3). Namely, butylparaben (110  $\mu$ M) reduced relative area by 55%; nilotinib (2  $\mu$ M) reduced relative area by 25%; and propylene glycol (850 mM) caused death. Overall, a significantly higher percentage of the positive exposures exhibited adverse impacts on EBs than the negative exposures ( $P = 0.013$ ; chi-square test).

No adverse effect was observed in the P19C5 gastrulation model in response to seven of the Daston's positive exposures, namely abacavir (80  $\mu$ M), artesunate (20 nM), dabigatran (7  $\mu$ M), ethylene glycol (57 mM), fingolimod (67 nM), glycolic acid (5 mM), and methanol (270 mM). Thus, we further examined whether EBs could be affected by these compounds at higher concentrations than those indicated in the Daston list. Artesunate, fingolimod, or methanol did not show any adverse effect on EB morphogenesis even at four times higher concentrations (Fig. 3.2), suggesting that the P19C5 gastrulation model is unable to detect developmental toxicity of these compounds at physiologically relevant concentrations. However, abacavir at four times higher concentration (320  $\mu$ M) reduced relative area by 25%; dabigatran at two times higher concentration (14  $\mu$ M) reduced relative area by 21%; ethylene glycol at four times higher concentration (228 mM) reduced relative area by 25%; and glycolic acid at four times higher concentration (20 mM) reduced relative area by 80% and reduced EDI by 70%. This suggests that the P19C5 gastrulation model may still be able to detect these four developmental toxicants, albeit with slightly less sensitivity than *in vivo*.

Ten of the Daston compounds evaluated (abacavir, ATRA, caffeine, dabigatran, ethylene glycol, glycolic acid, MEHP, methanol, nilotinib, and ramelteon) are categorized as both positive and negative exposures depending on concentrations. Such compounds are the most useful to rigorously validate the sensitivity and specificity of screening assays (Daston *et al.*, 2014). Accordingly, to the classification criteria described above, the P19C5 EB model was able to distinguish between positive and negative exposures for four of these compounds (ATRA, caffeine, MEHP, and ramelteon). If four times higher concentrations were also to be included for positive exposures, additional four compounds (abacavir, dabigatran, ethylene glycol, and glycolic acid) can also be differentially classified by the P19C5 EB model.

### 3.3.3 CYTOTOXICITY ON SOMATIC CELL LINES

Three of the negative exposures adversely impacted the P19C5 gastrulation model. Namely, butylparaben and nilotinib reduced EB size, whereas propylene glycol caused death of EBs. Thus, these negative exposures appeared to be detrimental to growth or survival of P19C5 EBs, raising the possibility that the gastrulation model may be overly sensitive to these compound exposures, leading to misclassifications. To test whether harmful effects of these exposures are unique to the P19C5 EBs, we performed cell viability assay using monolayer cultures of P19C5 cells as well as three types of somatic cell lines, NIH/3T3 (mouse embryonic fibroblast-derived), HEK293 (human embryonic kidney-derived), JEG3 (human trophoblast-derived). These cell lines were cultured for four days in the presence of butylparaben, nilotinib, or propylene glycol, and impact on cell proliferation or survival was scored. All cell lines were completely killed by propylene glycol at the negative exposure concentration (850 mM) (Fig. 3.4). Butylparaben at the negative exposure concentration (110  $\mu$ M) also consistently reduced relative cell numbers of all four cell lines (Fig. 3.4). Thus, these two negative exposures appeared to exhibit cytotoxic

effects on not only P19C5 cells but also on other cell lines. On the other hand, relative cell numbers were not consistently reduced by nilotinib in any of the cell lines at the negative exposure concentration (2  $\mu$ M) (Fig. 3.4).

### 3.4 - Discussion

The present study evaluated our *in vitro* gastrulation model of P19C5 stem cells in reference to the Daston list of compound exposures for developmental toxicity assays. Based on the morphogenetic effects on EBs, 10 out of 17 positive exposures were classified as adverse, whereas 14 out of 17 negative exposures were non-adverse. As shown in the previous study, valproic acid at 0.8 mM, another positive exposure in the Daston list, also causes adverse effect on P19C5 EB morphogenesis (Li and Marikawa, 2016). Thus, altogether, 25 out of 35 exposures in the Daston list (71.4%) were correctly classified by morphology-based assay using the P19C5 gastrulation model. When up to four times higher concentrations are included in assessment of the positive exposures, additional four exposures were also classified correctly, totaling 29 out of 35 (82.9%), although the validity of such an *ad hoc* criterion will require additional biological and toxicological justification.

To date, various types of non-animal alternatives have been explored as developmental toxicity screening assays, including those utilizing differentiation, migration, or metabolomics of mouse or human ES cells (Kuske *et al.*, 2012; Palmer *et al.*, 2013; Seiler and Spielmann, 2011; Theunissen *et al.*, 2012; Xing *et al.*, 2015), rodent whole embryo culture (Piersma *et al.*, 2004), micromass culture of limb bud mesenchyme (Pratten *et al.*, 2012), and non-mammalian model systems, namely zebrafish embryos (Sipes *et al.*, 2011). At present, it may not be fruitful to discuss effectiveness of the P19C5 gastrulation model in comparison with these assays, because no validation study has been reported in reference to the Daston list, to the best of our knowledge. Evaluations of other non-animal alternatives using the Daston list should provide solid framework for direct comparisons between different assays to reveal their strengths and weaknesses and help to assemble a proper battery of tests to screen a broad range of developmental toxicants.

In the present study, the effects of compound exposures on the P19C5 gastrulation model were evaluated solely based on the morphological features, namely the size and shape of EBs. However, the effects of compounds may also be evaluated through gene expression analyses, because P19C5 EBs exhibit distinct temporal and spatial gene expression patterns that are characteristic of early embryogenesis during axial elongation and patterning (Lau and Marikawa, 2014; Li and Marikawa, 2015; Marikawa *et al.*, 2009). Effects on gene expression are generally considered to be more sensitive endpoints when assessing developmental toxicity, and various types of gene expression analyses have been incorporated into stem cell-based *in vitro* assays, particularly ESTs, to augment their detection of developmental toxicants (Buesen *et al.*, 2009; Gao *et al.*, 2014; Panzica-Kelly *et al.*, 2013; Pennings *et al.*, 2011; Suzuki *et al.*, 2011; zur Nieden *et al.*, 2001). This notion applies to the response of P19C5 EBs to DAPT, a pharmacological inhibitor of Notch signaling, which markedly down-regulates expression of somitogenesis regulator genes, such as *Hes7*, *Lfng*, and *Nrarp*, without significantly altering EB size or shape (Li and Marikawa, 2015). However, the use of gene expression analyses to detect developmental



toxicants requires cautious interpretation, as the heightened sensitivity may increase the rate of false positive results. Mouse genetic studies, using the targeted gene knockout technology, have demonstrated in various cases that the loss of important developmental regulator genes fails to cause overt phenotypic abnormalities, due to existence of compensatory or redundant mechanisms (Barbaric *et al.*, 2007; Nowak *et al.*, 1997). Thus, alterations in gene expression may not correlate with the dramatic phenotypic effects (e.g., embryonic death or major fetal malformations) referenced in the Daston list.

Among the 34 exposures from the Daston list examined here, the negative exposure of ATRA (1.7 nM) was particularly unique in that it markedly increased relative area of EBs. Because relative EDI was concomitantly increased, ATRA at the low concentration appeared to promote growth and axial elongation of EBs. Although ATRA is teratogenic at excessively high concentrations, it also acts at low physiological concentrations as an endogenous regulator of cell proliferation and embryo body patterning (Clagett-Dame and Knutson, 2011). Inactivation of *Aldh1a2* gene, which encodes retinaldehyde dehydrogenase to synthesize ATRA, impairs growth and axial elongation in mouse embryos (Niederreither *et al.*, 1999). Likewise, treatment of P19C5 EBs with BMS493, a pharmacological antagonist of retinoic acid receptors, reduces relative area and EDI, suggesting that active retinoic acid signaling is essential for growth and axial elongation in the *in vitro* gastrulation model (Li and Marikawa, 2015). The source of retinoic acid or its precursors in the culture medium used in the present study was possibly bovine serum (Materials and Methods), and it may not have contained sufficient amount to support maximal growth and axial elongation of EBs. Thus, supplementation of a small amount of ATRA (as in the case for 1.7 nM exposure) would be beneficial for more robust EB development *in vitro*.

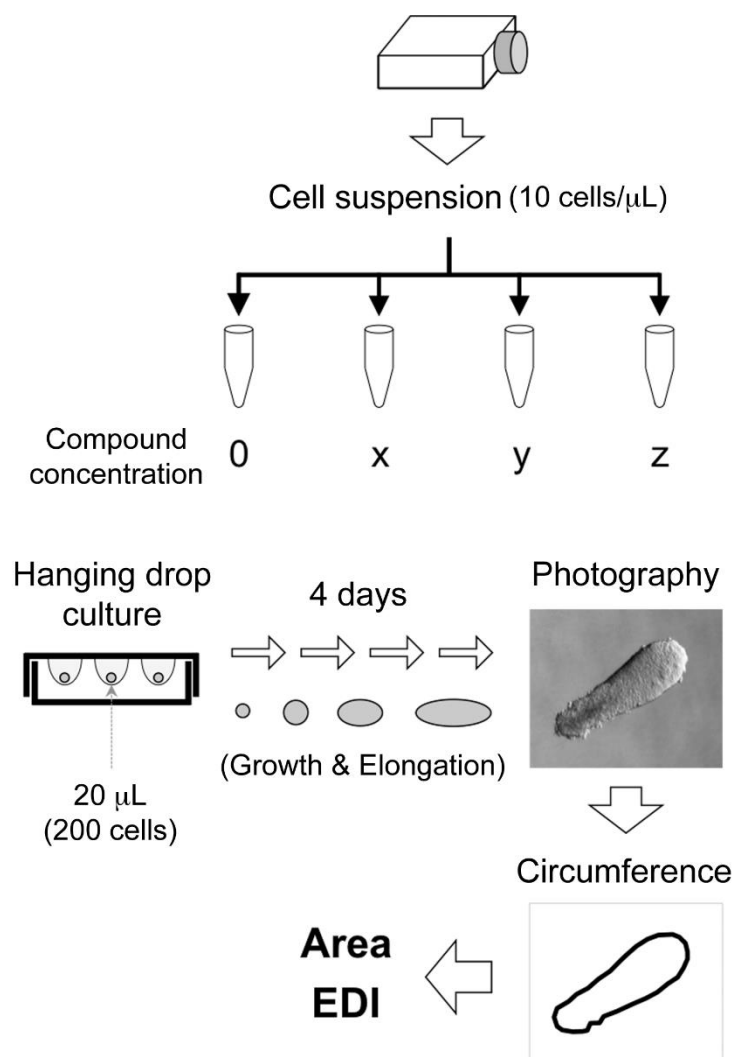
Artesunate, fingolimod and methanol caused no adverse effect on the P19C5 gastrulation model even at four times higher concentrations than those indicated as positive exposures in the Daston list. These “false negative” misclassifications may exemplify some of the limitations and weaknesses of the model. Artesunate, an anti-malaria medication, is rapidly converted to its active metabolite dihydroartemisinin, which exerts developmental toxicity by generating free radicals that damage embryonic erythroblasts (Clark *et al.*, 2008; Li *et al.*, 2009). Depletion of primitive embryonic erythroblasts likely accounts for the observed cardiovascular and skeletal abnormalities caused by artesunate treatment (Clark, 2009; White *et al.*, 2006; Clark *et al.*, 2008; White and Clark, 2008). The developmental toxicity of fingolimod, an immunomodulator, also manifests as abnormal vascular maturation and cardiac malformation, mediated by inhibition of the sphingosine-1 phosphate receptor (FDA, 2010; Schmid *et al.*, 2007). The P19C5 gastrulation model may not accurately identify such developmental toxicants that selectively disrupt hematologic processes, because *in vitro* development of EBs does not require a vascular system. On the other hand, the developmental toxicity of methanol may be mediated by its metabolic byproduct, formate (Andrews *et al.*, 1995). Compounds that become teratogenic only after chemical modifications by maternal metabolism are termed, “proteratogens” (Wells and Winn, 1996).

Detection of proteratogens is often considered a weakness of *in vitro* screens for developmental toxicants, because these assays usually lack the maternal metabolic system. Future studies will address whether the P19C5 model is susceptible to formate, and explore the incorporation of an exogenous metabolic activation system, such as co-culture with hepatocytes (Hettwer *et al.*, 2010; Oglesby *et al.*, 1986) or liver microsome fractions (Luijten *et al.*, 2008; Zhao *et al.*, 1993).

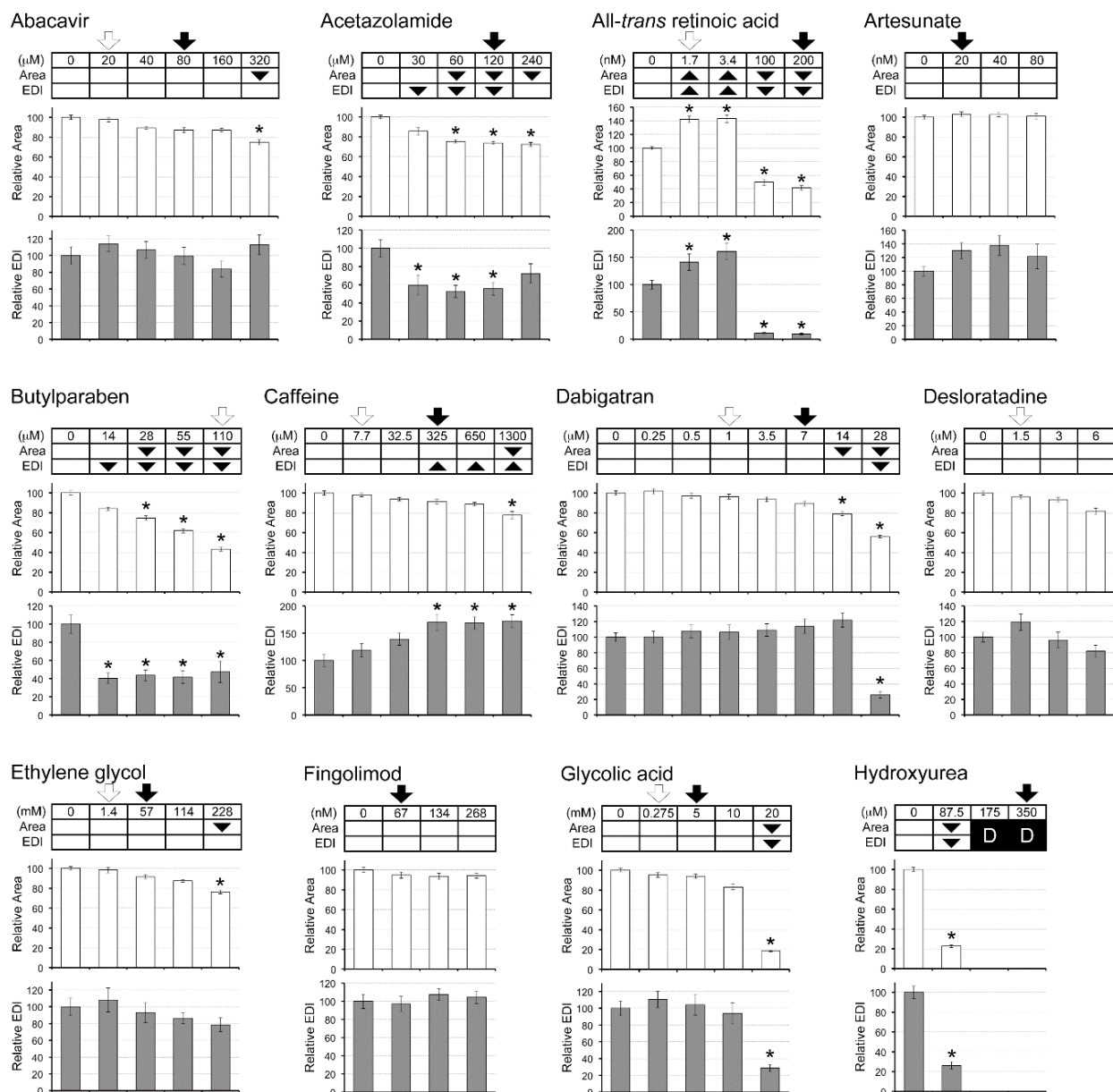
Practically speaking, misclassification of negative exposures (i.e., false positives) may be more detrimental for effectiveness of developmental toxicity assays, as compared to misclassification of positive exposures (i.e., false negatives). Unnecessary dismissal of safe compound exposures based on incorrect assay outcomes is disadvantageous for pharmaceutical development and burdensome for regulatory agencies (Waring *et al.*, 2015). In the case of the P19C5 gastrulation model, three negative exposures gave false positive outcomes: butylparaben at 110  $\mu\text{M}$  (reduction in area by 55%), nilotinib at 2  $\mu\text{M}$  (reduction in area by 25% and increase in EDI by 60%), and propylene glycol at 850 mM (death). The butylparaben and propylene glycol exposures diminished proliferation or survival of three somatic cell lines. Such general cytotoxicity may lead to misclassification of these exposures by other *in vitro* assays as well. It is unclear how such cytotoxic exposures are apparently harmless to developing embryos *in vivo* (Daston *et al.*, 2014). It is important to note that the compound concentrations indicated in the Daston list are  $C_{\text{max}}$ , and therefore developing embryos may not be exposed to such concentrations in a continuous manner, especially when compound clearance by the mother is rapid. Indeed, both butylparaben and propylene glycol have elimination half-lives of approximately four hours and undergo biotransformations *in vivo* (Aubert *et al.*, 2012; Morshed *et al.*, 1988). Additional pharmacokinetics information, namely area under the curve (AUC), may be more applicable to *in vitro* experimental conditions, and may help resolve this potential discrepancy. In contrast, the nilotinib negative exposure did not consistently diminish proliferation of the somatic cell lines, suggesting that adverse effect was specific to the P19C5 gastrulation model. Interestingly, nilotinib at the high concentration (28  $\mu\text{M}$ ; positive exposure in the Daston list), exhibited an unusual impact, i.e., P19C5 cells in hanging drops failed to aggregate while they were still alive and proliferating. Although spontaneous cell aggregation is an essential step in generating EBs, this process does not correspond to normal embryonic development. Thus, nilotinib may be altering cellular properties that are not embryologically relevant, and such non-physiological effects of nilotinib, even at lower concentrations, may have contributed to false positive misclassification of the compound exposure.

In the future, the exposure-based validation list (Daston *et al.*, 2014) will likely be expanded and refined as additional *in vivo* data on developmental toxicity and pharmacokinetics become available for various compounds. As previously suggested (Wise, 2016), it may be important to categorize positive exposures into distinct groups according to severity and nature of developmental toxicity. For example, the mechanism of an exposure that results in embryonic death is fundamentally different from an

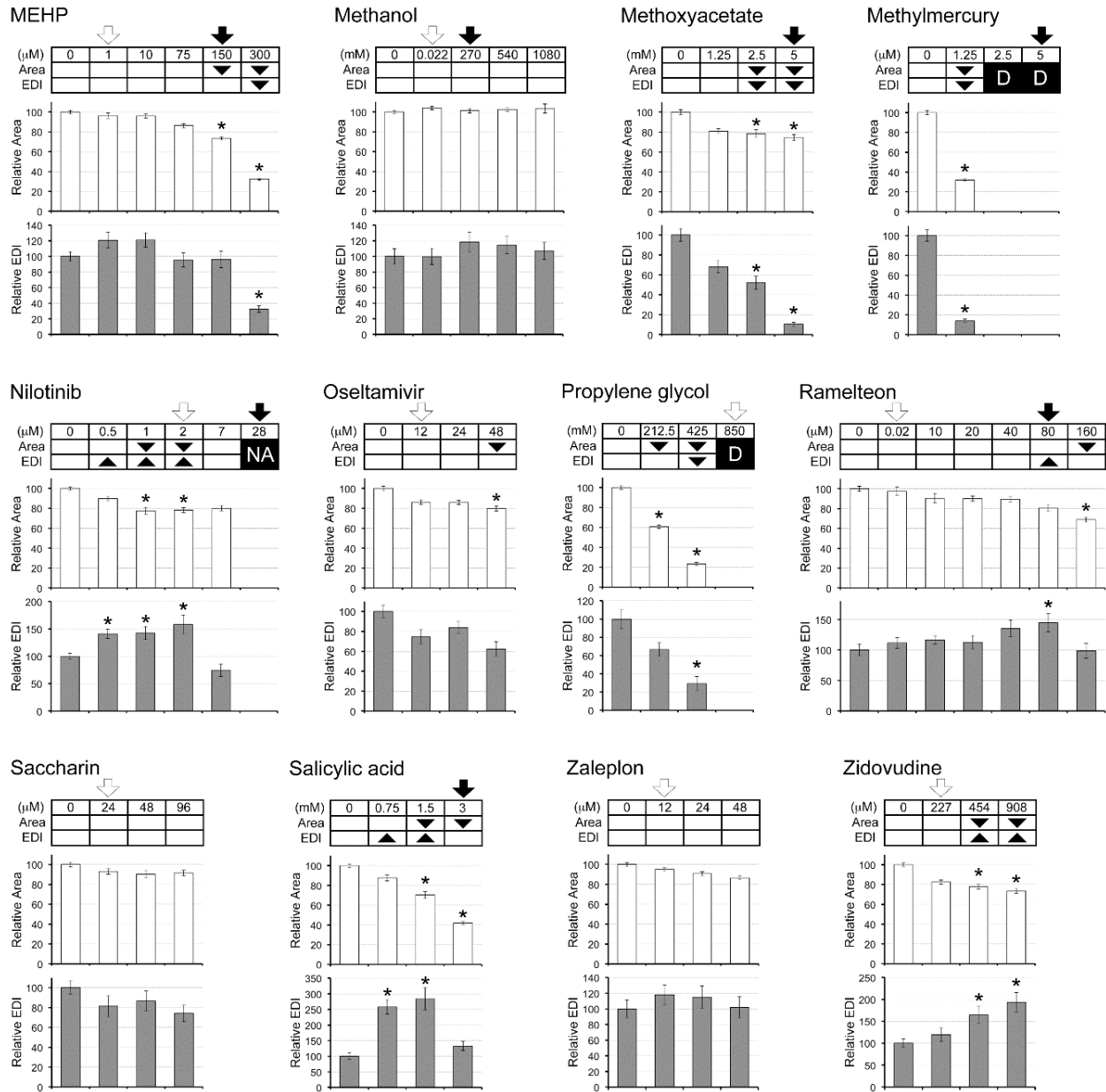
exposure that causes spina bifida or microcephaly. Should such differences in *in vivo* impact be reflected in the outcomes of *in vitro* assays? In the present study, the P19C5 gastrulation model exhibited various responses to the positive exposures, ranging from total degeneration to a reduction in axial elongation. The range of phenotypic responses warrants further investigations to examine whether these variations in EB integrity and morphology correlate with the nature and severity of *in vivo* effects observed after exposure to the corresponding developmental toxicants.



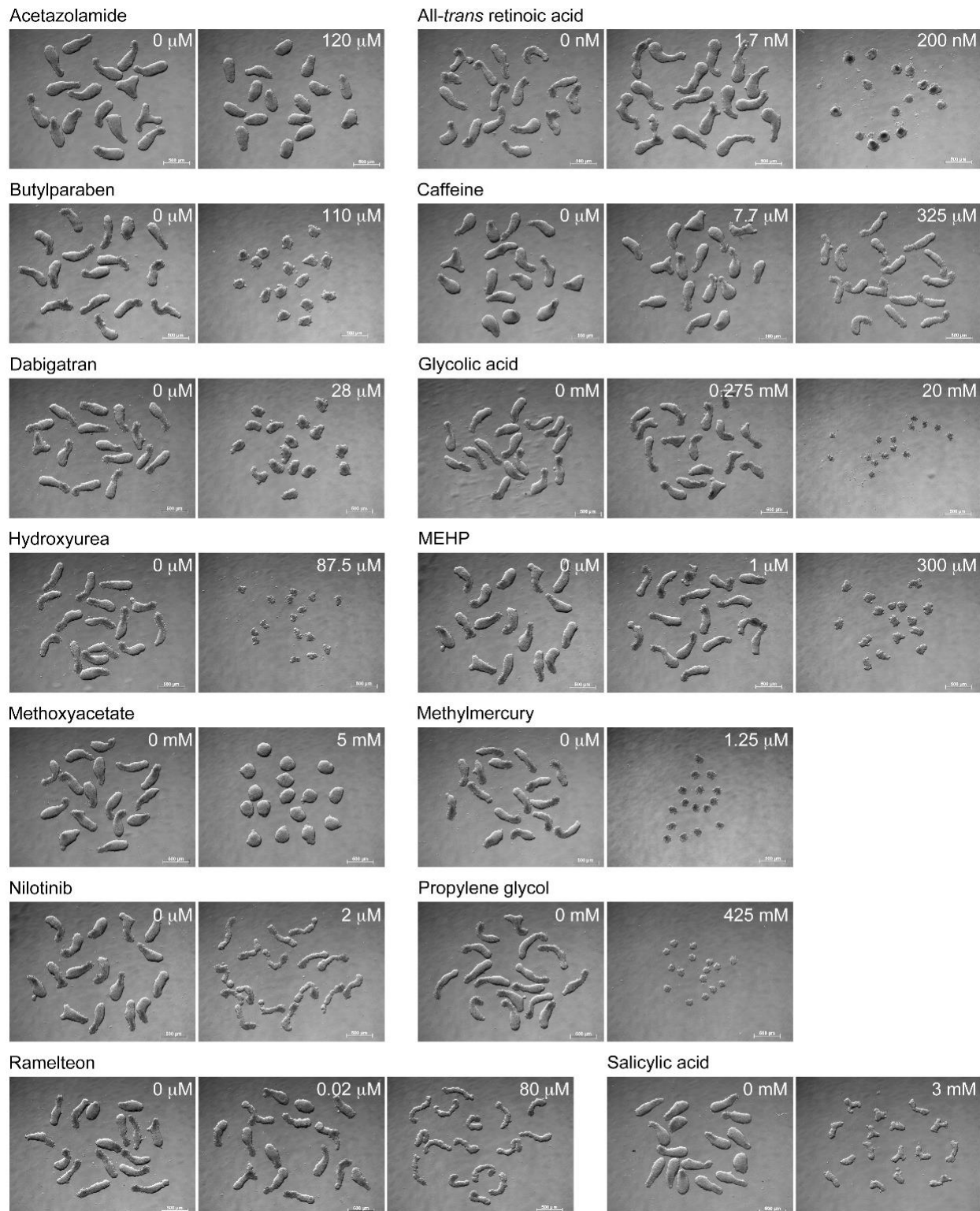
**Figure 3.1.** The experimental scheme to examine morphogenetic impact of the compound exposures. See Materials and Methods for details.



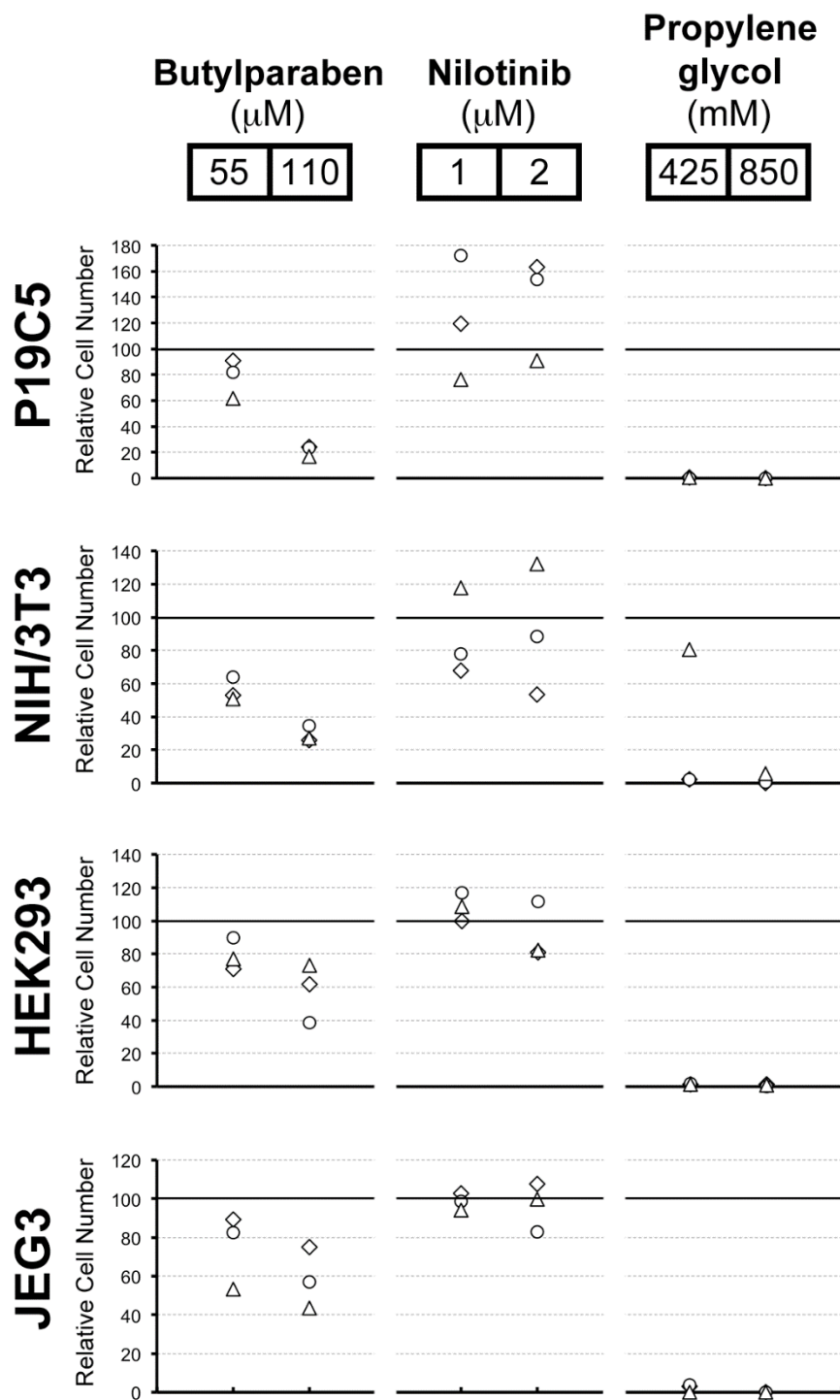
**Figure 3.2.** Impact of the Daston compounds on P19C5 EB morphogenesis. For each compound, concentrations tested are indicated in the top row of the table with a summary of observed morphogenetic impact on EB area and EDI, indicated with upward arrowheads (increase) and downward arrowheads (reduction). No area or EDI value is available when EBs were dead (D) or cells did not aggregate (NA). Arrows above the table indicate negative (white arrow) and positive (black arrow) exposures as cited in the Daston list. Column graphs below the summary tables show averages of relative area (white columns) and relative EDI (gray columns), for the corresponding compound concentrations indicated above. Error bars show 95% confidence intervals. Asterisks indicate adverse impacts, which are defined in the present study as a change in average area by >20% or a change in average EDI by >40% relative to controls.



**Figure 3.2. (Continued).** Impact of the Daston compounds on P19C5 EB morphogenesis. For each compound, concentrations tested are indicated in the top row of the table with a summary of observed morphogenetic impact on EB area and EDI, indicated with upward arrowheads (increase) and downward arrowheads (reduction). No area or EDI value is available when EBs were dead (D) or cells did not aggregate (NA). Arrows above the table indicate negative (white arrow) and positive (black arrow) exposures as cited in the Daston list. Column graphs below the summary tables show averages of relative area (white columns) and relative EDI (gray columns), for the corresponding compound concentrations indicated above. Error bars show 95% confidence intervals. Asterisks indicate adverse impacts, which are defined in the present study as a change in average area by >20% or a change in average EDI by >40% relative to controls.

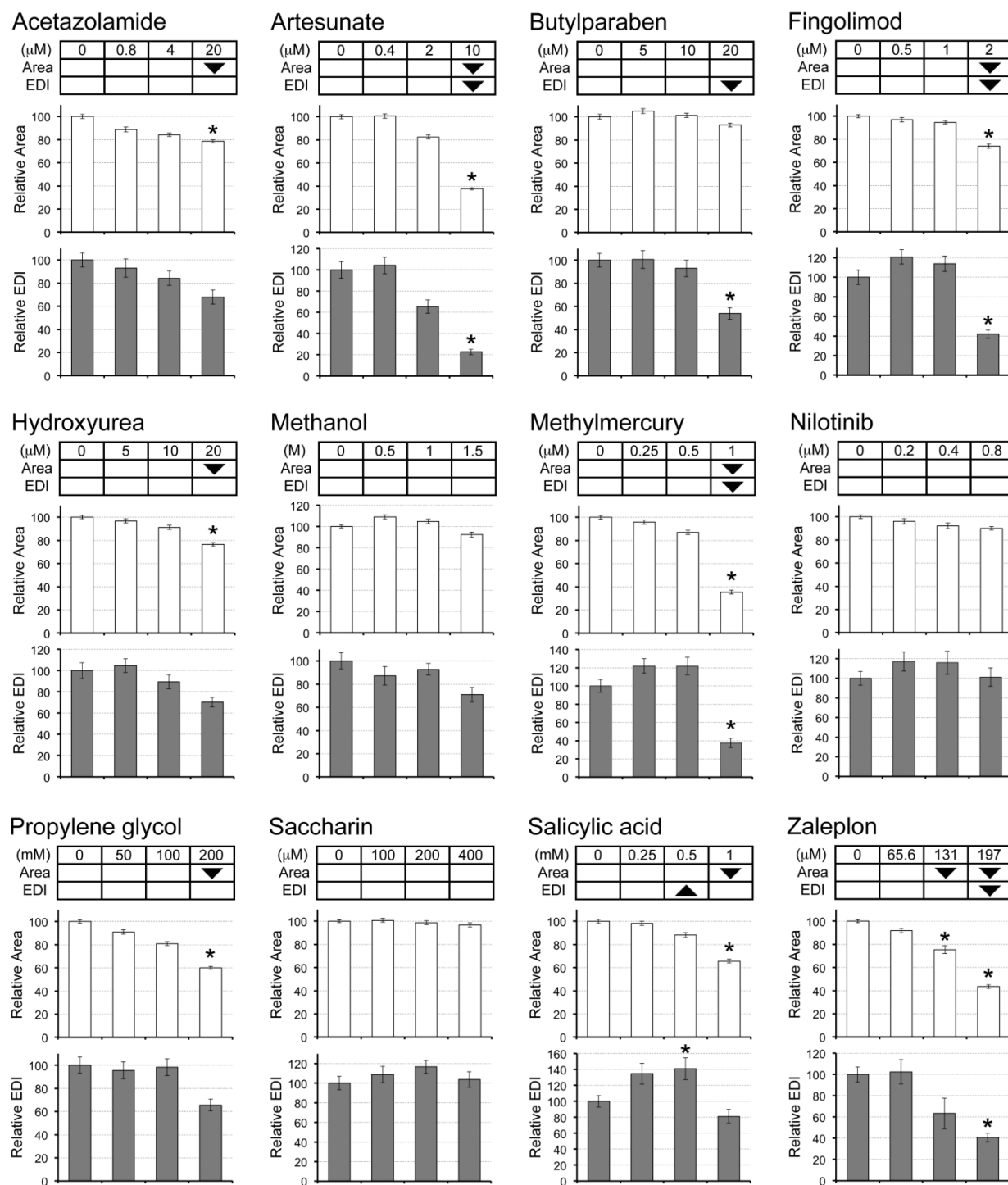


**Figure 3.3.** Representative images of P19C5 EBs. Not all compound exposures are shown (see text). Each set of images shows a control group of EBs (no test compound) and one or two compound-treated groups of EBs, made from the same cell suspension as the control. Scale bars = 500 μm.



**Figure 3.4.** Relative cell numbers of four cell lines, P19C5, NIH/3T3, HEK293 and JEG3, that are cultured as monolayer in the presence of butylparaben, nilotinib, or propylene glycol, at the negative exposure concentration indicated in the Daston list. Data points for three biological replicates are shown using different marker styles.





**Figure 3.5.** Extended ranges of treatment concentrations to identify the LOAELs of selected chemical exposures in the Daston list. Initial treatment ranges included the concentrations listed by Daston *et al.* and included 2- and 4-fold increases or decreases (Fig. 3.2). If the chemical showed no significant effects at the initial concentrations tested, then the range was increased to the LOAEL or the maximum possible treatment (e.g., artesunate, fingolimod, methanol, saccharin and zaleplon). Chemicals that had effects at concentrations lower than specified in the Daston list (e.g., acetazolamide, butylparaben, methylmercury, nilotinib, propylene glycol and salicylic acid) were treated with additional lower dilutions until the NOAEL was reached.

**Table 3.1.** Compounds used in the Daston validation study.

Compound	Vendor	Catalog Number	Stock
Abacavir	Sigma-Aldrich	SML0089	50 mM in water
Acetazolamide	Sigma-Aldrich	A6011	50 mM in DMSO
All-trans retinoic acid	Sigma-Aldrich	R2625	200 $\mu$ M in DMSO
Artesunate	Sigma-Aldrich	A3731	50 mM in DMSO
Butylparaben	Sigma-Aldrich	PHR1022	50 mM in DMSO
Caffeine	Sigma-Aldrich	C0750	1 M in water
Dabigatran	Santa Cruz Biotech.	sc-351724	20 mM in DMSO
Desloratadine	Sigma-Aldrich	D1069	1 mM in DMSO
Ethylene glycol	Sigma-Aldrich	324558	18 M (100%)
Fingolimod	Sigma-Aldrich	SML0700	20 mM in water
Glycolic acid	Sigma-Aldrich	G8284	1 M in water
Hydroxyurea	Sigma-Aldrich	H8627	100 mM in water
MEHP	Santa Cruz Biotech.	sc-396467	100 mM in DMSO
Methanol	Fisher Scientific	A412P-4	24.7 M (100%)
Methoxyacetate	Sigma-Aldrich	194557	10 mM in water
Methylmercury	Sigma-Aldrich	442534	10 mM in DMSO
Nilotinib	Santa Cruz Biotech.	sc-202245	50 mM in DMSO
Oseltamivir	Santa Cruz Biotech.	sc-208135	50 mM in water
Propylene glycol	Sigma-Aldrich	PHR1051	13.62 M (100%)
Ramelteon	Santa Cruz Biotech.	sc-219934	20 mM in DMSO
Saccharin	Sigma-Aldrich	240931	20 mM in water
Salicylic acid	Sigma-Aldrich	247588	1 M in DMSO
Zaleplon	Sigma-Aldrich	Z-004	3.28 mM in methanol
Zidovudine	Sigma-Aldrich	PHR1292	100 mM in DMSO

DMSO: Dimethylsulfoxide, MEHP: Mono(2-ethylhexyl) phthalate

**Table 3.2.** Adverse impacts of the Daston positive exposures on EB morphology.

Compound	Concentration	Adverse impact	Comments
Abacavir	80 µM	No	Area reduction at 4× higher conc. (320 µM)
Acetazolamide	121 µM	Yes (120 µM)	Area and EDI reduction
ATRA	200 nM	Yes	Area and EDI reduction
Artesunate	20 nM	No	
Caffeine	325 µM	Yes	EDI increase
Dabigatran	7 µM	No	Area reduction at 2× higher conc. (14 µM)
Ethylene glycol	57 mM	No	Area reduction at 4× higher conc. (228 mM)
Fingolimod	67 nM	No	
Glycolic acid	5 mM	No	Area and EDI reduction at 4× higher conc. (20 mM)
Hydroxyurea	350 µM	Yes	Dead
MEHP	146 µM	Yes (150 µM)	Area reduction
Methanol	270 mM	No	
Methoxyacetic acid	5 mM	Yes	Area and EDI reduction
Methylmercury	5 µM	Yes	Dead
Nilotinib	28 µM	Yes	No aggregation
Ramelteon	81 µM	Yes (80 µM)	EDI increase
Salicylic acid	3 mM	Yes	Area reduction

ATRA: All-*trans* retinoic acid, MEHP: Mono(2-ethylhexyl) phthalate

**Table 3.3.** Adverse impacts of the Daston negative exposures on EB morphology.

Compound	Concentration	Adverse impact	Comments
Abacavir	18 µM	No (20 µM)	
ATRA	1.7 nM	No	EDI increase accompanied by area increase (see text)
Butylparaben	110 µM	Yes	Area and EDI reduction
Caffeine	7.7 µM	No	
Dabigatran	1 µM	No	
Desloratadine	1.5 µM	No	Concentration listed as 1.5 mM in Daston <i>et al.</i> 2014.
Ethylene glycol	1.4 mM	No	
Glycolic acid	275 µM	No	
Methanol	22 µM	No	
MEHP	1 µM	No	
Nilotinib	2 µM	Yes	Area reduction, EDI increase
Oseltamivir	12 µM	No	
Propylene glycol	850 mM	Yes	Dead
Ramelteon	19 nM	No	
Saccharin	24 µM	No	
Zaleplon	12 µM	No	
Zidovudine	227 µM	No	

ATRA: All-*trans* retinoic acid, MEHP: Mono(2-ethylhexyl) phthalate

## CHAPTER 4. FLUOXETINE INHIBITS CANONICAL WNT SIGNALING TO IMPAIR EMBRYOID BODY MORPHOGENESIS: POTENTIAL TERATOGENIC MECHANISMS OF A COMMONLY USED ANTIDEPRESSANT

### 4.1 - Introduction

Many chemical compounds have the potential to disrupt sensitive processes of embryonic development, resulting in miscarriage or birth defects. For this reason, pregnant women are advised to avoid using any unnecessary medications, supplements or recreational drugs. However, some women with pre-existing medical conditions, such as depression, must continue taking medications during their pregnancies to maintain maternal physical and mental well-being. Depression affects 8 to 16% of women of reproductive age in the United States (Farr *et al.*, 2010; Ko *et al.*, 2012; Willet *et al.*, 2012). Because untreated depression during pregnancy is associated with a range of adverse pregnancy outcomes, such as miscarriage, preterm birth, and lower birth weights (Udechuku *et al.*, 2010), antidepressants are often prescribed for use during pregnancy. Unfortunately, some epidemiologic studies have suggested a higher incidence of birth defects in children exposed to antidepressants during embryonic development. To minimize the risk of preventable birth defects in human pregnancies, it is important to know which antidepressants are most likely to be teratogenic and which are least likely to cause harm if used during pregnancy. This understanding can only be obtained through continued investigations on the developmental toxicity of antidepressants using retrospective human epidemiologic studies as well as more prospective and mechanistic analyses using animal and *in vitro* models. With such information, clinicians can differentiate various types of antidepressants based on their relative safety for use during pregnancy and prescribe the safest medications to the patients.

Fluoxetine (a.k.a. Prozac) is one of the most commonly prescribed antidepressants of the selective serotonin reuptake inhibitor (SSRI) class. By inhibiting the serotonin transporter (SERT; encoded by the *SLC6A4* gene), fluoxetine and other SSRIs increase the synaptic levels of serotonin (5-hydroxytryptamine or 5-HT), a monoamine neurotransmitter that helps to regulate mood. Several human epidemiologic studies have identified significant associations between maternal use of fluoxetine during pregnancy and an increased incidence of specific congenital anomalies, such as cardiac defects, pulmonary hypertension, gastroschisis and omphalocele (Ellfolk and Malm, 2010; Reefhuis *et al.*, 2015; Wemakor *et al.*, 2015)). The mechanism linking fluoxetine to these birth defects is unclear, but there are three primary possibilities: fluoxetine may inhibit SERT in the fetus, inhibit SERT in the mother, or affect an off-target molecule unrelated to serotonin signaling altogether. In the mouse embryo, SERT is expressed in the heart and liver as early as the E10.5 stage (Narboux-Nême *et al.*, 2008). Fluoxetine may inhibit SERT in these embryonic tissues to directly cause organ malformations. Alternatively, the developmental toxicity of fluoxetine may be indirect if it acts on SERT in the pregnant mother, altering maternal physiology (i.e., decreasing placental blood flow) and resulting in adverse effects on embryonic

development. For example, fluoxetine causes severe anorexia in rats, leading to maternal starvation with secondary adverse reproductive outcomes (Sloot *et al.*, 2009; Wong *et al.*, 1988). Finally, fluoxetine may target non-SERT molecules that play essential roles in embryogenesis. *In vivo* studies with humans or model animals are often too complex to distinguish which of these possibilities is responsible for the teratogenic action of fluoxetine. On the other hand, *in vitro* tests are devoid of maternal factors and are more amenable to experimental manipulations and molecular analyses. Thus, *in vitro* models may provide valuable mechanistic insight into the developmental toxicity of fluoxetine.

One such *in vitro* model uses stem cell-derived embryoid bodies capable of spontaneous axial morphogenesis to study the effects of developmental toxicants. These EBs are created from P19C5 mouse stem cells, which possess developmental properties similar to the epiblast, the pluripotent embryonic precursor of the fetal body (Lau and Marikawa, 2014). P19C5 cells can be aggregated in a hanging drop of culture medium to differentiate as an embryoid body (EB), which transforms from a spherical mass to an elongated structure during four days of culture. This morphological transformation is an *in vitro* recreation of gastrulation, the morphogenetic process of body patterning and elongation along the cranial-caudal embryonic axis. The developmental toxicity of chemical agents can be assessed based on their adverse impact on the *in vitro* morphogenesis of P19C5 EBs, a model that was recently validated using the Daston list of reference chemical exposures (Daston *et al.*, 2014; Warkus and Marikawa, 2017). The Daston list consists of 39 exposures, i.e., *in vivo* concentrations of specific compounds that cause adverse effects on embryos or lack thereof (Daston *et al.*, 2014). Growth and morphogenesis of P19C5 EBs were significantly altered by the adverse exposures in the Daston list, but not by the non-adverse exposures, with a total concordance of 71.4 to 82.9% (Warkus and Marikawa, 2017). Thus, the morphogenesis-based P19C5 EB assay can serve as an effective *in vitro* alternative to evaluate developmental toxicity of compounds at physiologically-relevant concentrations.

The P19C5 EB morphogenesis model is also a useful tool to identify the molecular mechanisms of developmental toxicity. During the four days of culture, EBs exhibit dynamic gene expression patterns in a distinct temporal pattern that is similar to the time course of gene regulation in the *in vivo* embryo. For example, pluripotency maintenance genes (e.g., *Pou5f1* and *Nanog*) are down-regulated by Day 1 of EB culture, whereas genes associated with the initial step of gastrulation (e.g., *Brachyury*, *Mixl1*, and *Cdx1*) show a strong peak expression on Day 1. In contrast, the genes involved in caudal development and somite formation have peak expression levels on Day 2 (e.g., *Wnt3a*, *Tbx6*, *Notch1*, and *Hes7*) or Day 3 (e.g., *Meox1* and *Mesp2*) (Lau and Marikawa, 2014; Li and Marikawa, 2015, 2016) (Table 4.1). The intensity and temporal patterns of these gene expressions are regulated by specific molecular signals, namely Wnt, Nodal, Fgf, Bmp, Notch, and retinoic acid, which was demonstrated using pharmacological inhibitors against each of these signaling pathways (Li and Marikawa, 2015). Disruption of a particular signal in EBs by the inhibitor causes distinct alterations in the gene expression patterns, yielding

“disruption profiles” that are specific to each signaling pathway. Comparison of these disruption profiles to the gene expression patterns of EBs exposed to a teratogen can provide mechanistic insight to identify which molecular signals are impacted by the teratogenic exposure. This strategy was demonstrated to be effective in a recent study, which revealed that valproic acid (a well-known teratogen) adversely affects morphogenesis by enhancing retinoic acid signaling (Li and Marikawa, 2016).

In the present study, we used the P19C5 EB system to investigate the mechanism of fluoxetine's effects on embryonic morphogenesis, particularly concerning its adverse concentration levels, impact on gene expression profiles, and molecular pathways that are independent of SERT inhibition.

## 4.2 - Materials and methods

### 4.2.1 COMPOUNDS

All chemical compounds used in the present study were commercially obtained, and their details are described in Table 2.

### 4.2.2 CELL CULTURE AND EMBRYOID BODY GENERATION

P19C5 cells were cultured and induced for *in vitro* morphogenesis, according to the methods described previously (Lau and Marikawa, 2014; Warkus and Marikawa, 2017). Briefly, cells were maintained in the P19 culture medium (Minimum Essential Medium Alpha, nucleosides, GlutaMAX Supplement [LifeTechnologies, Carlsbad, California], 2.5% fetal bovine serum, 7.5% newborn calf serum, 50 units/mL penicillin and 50 µg/mL streptomycin). For embryoid body (EB) generation, cells were dissociated with TrypLE Express (LifeTechnologies) and suspended at a density of 10 cells/µL in the culture medium containing a final concentration of 1% (v/v) dimethyl sulfoxide (DMSO) with or without a test compound. Droplets (20 µL each) of cell suspension were deposited on the inner surface of Petri dish lids and were cultured for up to 4 days in an incubator with 4.5% CO<sub>2</sub> at 37°C in humidified air. EBs were monitored daily for survival and overall integrity before being harvested for morphometric and gene expression analyses (Fig. 4.1).

### 4.2.3 MORPHOMETRIC ANALYSES

Images of EBs were obtained after 4 days of hanging drop culture (Day 4) and opened in the ImageJ program (<http://rsb.info.nih.gov/ij>) for morphometric analyses, as previously described (Warkus and Marikawa, 2017). Briefly, the outline (perimeter) of individual EBs was automatically identified using a macro and measured to obtain morphometric parameters, such as area and circularity (Fig. 4.10). Area was used to estimate the overall size of EBs, whereas the elongation distortion index (EDI), which is calculated from circularity ( $EDI = 1/circularity - 1$ ), was used to assess the extent of EB axial elongation. Area and EDI were normalized by the average values of control EBs in each set of experiments, and were reported as relative area and relative EDI, respectively, expressed in percentages (the average values of control are 100%). Three sets of experiments were conducted for each treatment condition as biological replicates, and relative area and relative EDI were compiled and presented as mean ± 95% confidence intervals. Adverse morphogenetic effects were defined as a reduction in relative area by >20% and/or a change in relative EDI by >40%, which are thresholds based on the criteria established in the previous validation study (Warkus and Marikawa, 2017).



#### 4.2.4 REVERSE TRANSCRIPTION AND POLYMERASE CHAIN REACTION (RT-PCR) ASSAYS

Total RNA was extracted from cell suspension prior to aggregation (Day 0) and from EBs on Days 1, 2, 3, and 4, using TRI reagent (LifeTechnologies) and Direct-zol RNA MiniPrep kit (Zymo Research, Irvine, CA), and processed for cDNA synthesis using M-MLV Reverse Transcriptase (Promega, Madison, WI) and oligo-dT (18) primer. The preparation of cDNA from the mouse whole embryo was described previously (Li and Marikawa, 2015). Standard RT-PCR was used to detect mRNAs of *Slc6a4* (encoding the serotonin transporter) and *Slc6a2* (encoding the norepinephrine transporter), using JumpStart REDTaq DNA Polymerase (Sigma-Aldrich) with the following conditions: initial denaturation at 94°C for 5 min; 35 cycles of 94°C for 15 sec, 60°C for 20 sec, and 72°C for 40 sec; and a final extension at 72°C for 5 min. Amplified products were resolved in a 1% agarose gel and visualized by staining with ethidium bromide. Note that primers for *Slc6a4* (F: 5'-GTT CTG CAG CGA CGT GAA GGA AAT-3', R: 5'- GCT TAG AGG GGA GGA GTC AAG GTG-3') and *Slc6a2* (F: 5'- GTG GTG GTC AGC ATC ATC AAC TTC-3', R: 5'- AAC CAG CGT CAC GGA ATC ATT AGT-3') were designed as intron-spanning, so that PCR products amplified from cDNA could be distinguished from those from genomic DNA based on the product size (416 and 420 bp for *Slc6a4* and *Slc6a2* mRNA, respectively).

Quantitative RT-PCR was performed using the CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) as follows: initial denaturation at 94°C for 5 min, followed by up to 45 cycles of 94°C for 15 sec, 60°C for 20 sec, and 72°C for 40 sec. Data files were opened in CFX Manager software (Bio-Rad) and Ct values were transferred to the Excel program for further analyses. *Actb*, which encodes  $\beta$ -Actin, was used as a constitutively-expressed housekeeping gene to standardize the expression levels of the other genes. Relative expression level was calculated as a percentage relative to the total expression level of control EBs (a sum of Days 0, 1, 2, 3 and 4 values) in each set of experiments. Three sets of experiments were conducted for each treatment condition as biological replicates, and relative expression levels were compiled and presented as mean  $\pm$  standard deviation.

#### 4.2.5 LUCIFERASE REPORTER ASSAYS

The plasmids used for luciferase reporter assays were obtained commercially or from other researchers. TOPFLASH (Upstate, Charlottesville, VA) consists of the firefly luciferase gene controlled by the transcriptional promoter containing multiple TCF-binding sites and serves as a reporter for active canonical Wnt signaling. FOPFLASH (Upstate), used as a negative control for TOPFLASH, contains inactivating mutations in the TCF-binding sites. pRL-TK (Promega) encodes the *Renilla* luciferase gene under the control of a constitutive promoter (thymidine kinase promoter) and was used to normalize the luciferase activities of TOPFLASH and FOPFLASH. L1CBD-G4DBD encodes a chimeric protein

containing the  $\beta$ -Catenin-binding domain of LEF1 and the DNA-binding domain of GAL4. Like TOPFLASH, L1CBD-G4BD serves as a monitoring tool for active canonical Wnt signaling, as previously described (Tamashiro *et al.*, 2008). Binding of  $\beta$ -Catenin to L1CBD-G4DBD forms a transcriptional activator to turn on the expression of the firefly luciferase from pG5-Luc, which contains multiple GAL4-binding sites (Promega). G4DBD, which is GAL4-binding domain without a  $\beta$ -Catenin domain, was used as a control for L1CBD-G4DBD (Tamashiro *et al.*, 2008). ARE-Luc (a gift from Dr. Malcom Whitman; Harvard Medical School) consists of the firefly luciferase controlled by the activin response elements and serves as a reporter for active Nodal/Activin signaling (Weisberg *et al.*, 1998). P19C5 cells in monolayer were transfected with the plasmids using Lipofectamine 2000 (LifeTechnologies). After 24 hours, cells were lysed and examined for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega) with Gene Light 55 Luminometer (Microtech, Chiba, Japan). All experiments were repeated independently three times. The results are shown as mean  $\pm$  standard deviation.

#### 4.2.6 CELL VIABILITY ASSAY

Effect of chemical treatment on cell proliferation and viability was determined with the CellTiter-Glo Luminescent Cell Viability Assay System (Promega), as described previously (Warkus *et al.*, 2016). Briefly, P19C5 cells were seeded in 96-well plates at a density of 100 cells/well in 100  $\mu$ L of culture medium supplemented with 1% DMSO containing the test compound or vehicle only as a control. After 4 days of culture, cells were treated with CellTiter-Glo Reagent. The resulting luminescence was measured using Gene Light 55 Luminometer as a readout of ATP amount, which serves as a quantitative proxy for the number of metabolically active cells. The intensity of the luminescence was normalized to the control level (vehicle only) in each set of experiments and reported as relative light units. All experiments were repeated independently three times, and the results are shown as mean  $\pm$  standard deviation.

#### 4.2.7 STATISTICS

Statistical differences were assessed by two-sample t-test. For EB morphogenesis, relative area and relative EDI were compared between two groups of EBs, typically compound-treated EBs and vehicle-treated control EBs ( $n = 45\text{--}48$  for each group, compiled from 3 sets of independent experiments), unless otherwise stated. Differences in average values were deemed significant when P values were  $< 0.01$ . For luciferase reporter assays, normalized luciferase activities, i.e., ratios of the firefly luciferase to the *Renilla* luciferase signals, from three independent experiments were compared between compound-treated and control (vehicle only) groups. For the cell viability assay, relative light units from three independent experiments were compared between compound-treated and control cells.

## 4.3 - Results

### 4.3.1 FLUOXETINE IMPAIRS MORPHOGENESIS OF P19C5 EMBRYOID BODIES IN A DOSE-DEPENDENT MANNER

The teratogenic potential of fluoxetine was evaluated using the P19C5 embryoid body (EB) morphogenesis system (Fig. 4.1). P19C5 cells were aggregated in hanging drops of culture medium containing various concentrations of fluoxetine, and the morphology of the resulting EBs was examined after four days of culture. We focused on a concentration range between 2 and 10  $\mu\text{M}$ , because our pilot study suggested that fluoxetine at 1  $\mu\text{M}$  or less had no morphological impact whereas fluoxetine concentrations higher than 10  $\mu\text{M}$  were detrimental to the survival of EBs. Treatment effects were quantified using two morphometric parameters, the two-dimensional area of EBs, which is a proxy for EB size, and the elongation distortion index (EDI), which was used to evaluate the extent of axial elongation. Following the previous validation study, morphologic impacts were only considered to be adverse if drug treatment caused a reduction in relative area by >20% or a reduction in relative EDI by >40% compared to control EBs (Warkus and Marikawa, 2017). EB morphogenesis was adversely affected by fluoxetine treatment at the exposures levels of 6  $\mu\text{M}$  and higher, which reduced both relative area and relative EDI in a dose-dependent manner (Fig. 4.2A, B).

We then evaluated the effects of R-fluoxetine, S-fluoxetine, and norfluoxetine on EB morphogenesis. Fluoxetine is a racemic mixture of R- and S-enantiomers. R-fluoxetine and S-fluoxetine are similar in their potency as serotonin reuptake inhibitors (SSRIs), although the latter (S) is slightly more potent than the former (R) (Wong *et al.*, 1995). Norfluoxetine, the major metabolite of fluoxetine, is a potent SSRI with a longer half-life in the circulating plasma (Hiemke and Härtter, 2000). The morphogenetic effects of these compounds were tested at 2, 6, 10, and 20  $\mu\text{M}$ , to match the experiments done with fluoxetine. All the compounds adversely affected EB morphogenesis at 6  $\mu\text{M}$  and higher, although S-fluoxetine appeared slightly less potent than the others with respect to the extent of reduction in area and EDI (Fig. 4.2C, D). Thus, the morphogenesis of EBs was susceptible to exposures to fluoxetine, its enantiomers, and metabolite with largely similar dose-dependency.

### 4.3.2 THE MORPHOGENETIC EFFECTS OF FLUOXETINE ARE NOT DUE TO INHIBITION OF THE SEROTONIN TRANSPORTER

We assessed whether the action as an SSRI is responsible for the adverse morphogenetic effects of fluoxetine. First, expression of the *Slc6a4* gene was examined in P19C5 EBs, as it encodes the serotonin transporter (SERT), the main target of SSRI action. For comparison, expression of the *Slc4a2* gene, which encodes the norepinephrine transporter (NET) was also examined, as both SERT and NET

are inhibited by antidepressants of the serotonin and norepinephrine reuptake inhibitor (SNRI) class, such as venlafaxine. Both *Slc6a4* and *Slc6a2* mRNAs were robustly expressed in the mouse embryo at the stage E10.5. Conversely, in P19C5 EBs (as a mixture of Days 0, 1, 2, 3, and 4), only *Slc6a2* mRNA was robustly expressed (Fig. 4.3A). Quantitative RT-PCR analysis showed that *Slc6a2* mRNA was highly expressed only at Day 0 (i.e., before cell aggregation), and that the expression of *Slc6a4* mRNA was very low—less than 5% of the whole embryo level—at all stages of EBs (Fig. 4.3B).

Although the level was very low, the expression of SERT may be sufficient to cause adverse effects in EBs if inhibited by fluoxetine. Thus, we evaluated the morphogenetic impact of other inhibitors of SERT, specifically citalopram (SSRI) and venlafaxine (SNRI). EB morphology was assessed after four days of treatment with either citalopram or venlafaxine at 0, 2, 6, 10 or 20  $\mu$ M. Neither citalopram nor venlafaxine exhibited adverse effects even at the highest concentrations tested (Fig. 4.3C, D). Since SSRIs elevate the levels of extracellular 5-HT, we also tested the impact of excessive 5-HT on EB morphogenesis. None of the concentrations tested (ranging from 10  $\mu$ M to 100  $\mu$ M) caused adverse morphogenetic effects (Fig. 4.3E, F). These results indicate that the SSRI action is not responsible for the adverse morphogenetic effects of fluoxetine and suggest that fluoxetine may target molecular mechanisms unrelated to SERT when it interferes with embryonic morphogenesis.

#### **4.3.3 FLUOXETINE ALTERS EXPRESSION PATTERNS OF VARIOUS DEVELOPMENTAL REGULATOR GENES**

To gain mechanistic insight into the morphogenetic effects of fluoxetine, we examined how fluoxetine treatment alters expression patterns of key developmental regulator genes in EBs. EBs were treated with fluoxetine at the concentrations of 0  $\mu$ M (control), 2  $\mu$ M and 6  $\mu$ M, and harvested at Days 0, 1, 2, 3, and 4 for gene expression analyses by quantitative RT-PCR (Fig. 4.1). These concentrations were selected because 2  $\mu$ M had no morphogenetic impact, whereas 6  $\mu$ M was the lowest concentration that altered both area and EDI (Fig. 4.2B). The developmental regulators examined are transcription factors and signaling molecules that are crucial for embryo patterning and germ layer formation, and their characteristics are summarized in Table 1.

Fluoxetine at 2  $\mu$ M did not significantly alter the transcript levels of any of the developmental regulators during EB development (Fig. 4.4A). However, fluoxetine at 6  $\mu$ M changed the expression patterns of a number of genes relative to the patterns in vehicle-treated controls. In control EBs, the levels of the pluripotency maintenance genes (i.e., *Pou5f1* and *Nanog*) were markedly down-regulated by Day 1, whereas the genes involved in the initial phase of gastrulation (i.e., *Brachyury*, *Mixl1* and *Cdx1*) were strongly up-regulated, indicating a swift initiation of the differentiation program within one day of aggregation culture. In contrast, EBs treated with fluoxetine (6  $\mu$ M) exhibited a delay in differentiation,

such that the pluripotency genes continued to be highly expressed at Day 1 and the up-regulation of the initial gastrulation genes was less robust than control (Fig. 4.4A). A delay in differentiation was also evident in the expression patterns of other genes, such as the regulators of caudal patterning and somite formation (i.e., *Wnt3a*, *Tbx6*, *Hes7*, and *Meox1*). In these genes, the up-regulation towards the peak expression (on Day 2 for *Wnt3a*, *Tbx6*, and *Hes7*, and on Day 3 for *Meox1*) was diminished in fluoxetine-treated EBs compared to control EBs (Fig. 4.4A). This delayed differentiation seemed specific to the mesodermal lineage since genes expressed in the neural lineage (i.e., *Sox2*, *Otx2*, *Hoxc6*, and *Pax3*) were either elevated or largely unaffected by fluoxetine treatment (Fig. 4.4A). These results suggest that fluoxetine treatment (6  $\mu$ M) is detrimental for mesodermal differentiation but is neutral or even permissive for neural differentiation in P19C5 EBs.

Several major developmental signaling pathways, namely Wnt, Nodal, Fgf, Bmp, and retinoic acid (RA), regulate germ layer differentiation and body patterning during embryogenesis. These major pathways also control morphogenesis and gene expression in P19C5 EBs (Li and Marikawa, 2015). To assess whether fluoxetine alters these major developmental signals to cause adverse effects, we compared the gene expression profiles in fluoxetine-treated EBs with the “disruption profiles” of these signaling pathways. The previous study (Li and Marikawa, 2015) showed that the expression patterns of *Pou5f1*, *Nanog*, and *Brachyury* were differentially affected by inhibition of each of the major signaling pathways at Day 1 and Day 2, and, therefore, may be used to identify which pathway is disrupted (Fig. 4.4B). The impact of fluoxetine (6  $\mu$ M) was most similar to that of XAV939 (5  $\mu$ M), an inhibitor of the canonical Wnt signaling pathway (Fig. 4.4B). However, two other signaling inhibitors (SB431542 [Nodal inhibitor] and PD173074 [Fgf inhibitor]) also caused a decrease in the Day 1 expression of the early mesoderm gene *Brachyury*, similarly to the impact of fluoxetine. To assess whether these three signaling pathways—Wnt, Nodal, and Fgf—are linked to the molecular impact of fluoxetine, we examined the target genes of these pathways: *Sp5* for canonical Wnt signaling, *Nodal* for Nodal signaling, and *Spry2* for Fgf signaling. Fluoxetine treatment caused a reduction in the expression of *Sp5* at Day 1 but did not reduce *Nodal* or *Spry2* expression (Fig. 4.4A), suggesting that fluoxetine at 6  $\mu$ M may inhibit canonical Wnt signaling during EB development to alter EB morphogenesis.

#### **4.3.4 FLUOXETINE INHIBITS CANONICAL WNT SIGNALING IN P19C5 CELLS**

We tested whether fluoxetine inhibits canonical Wnt signaling using two luciferase reporter assays, both of which measure the transcriptional activation mediated by  $\beta$ -Catenin (CTNNB1) as a readout of Wnt signaling activity (Fig. 4.5A). One assay employs TOPFLASH (Korinek *et al.*, 1997), and the other utilizes L1CBD-G4DBD (Tamashiro *et al.*, 2008). P19C5 cells in monolayer were transfected with the reporter constructs and treated with fluoxetine at various concentrations for 24 hours. The TOPFLASH signal was significantly reduced by fluoxetine at 6  $\mu$ M or higher in a dose-dependent manner,

whereas the negative control FOPFLASH (containing mutations in the TCF-binding sites) yielded low signals regardless of fluoxetine concentration (Fig. 4.5B). Similarly, the L1CBD-G4DBD-dependent luciferase activity was significantly reduced by fluoxetine at 6  $\mu$ M or higher (Fig. 4.5C). In contrast, the activity of ARE-Luc, a reporter construct for Nodal/Activin signaling (Weisberg *et al.*, 1998), was not significantly reduced by fluoxetine or XAV939, whereas it was robustly repressed by SB431542 (Fig. 4.5D). These results indicate that fluoxetine inhibits canonical Wnt signaling at the same concentrations that adversely affected EB morphogenesis.

The impact of the fluoxetine enantiomers and the other SERT inhibitors was also assessed using the luciferase assays to further clarify the relationship between Wnt inhibition and morphogenetic effects. Both R-fluoxetine and S-fluoxetine reduced the TOPFLASH signal in a dose-dependent manner (Fig. 4.5E). S-fluoxetine appeared to be slightly less potent than R-fluoxetine, which corresponded to the observed impact on EB morphogenesis (Fig. 4.2C). The TOPFLASH signal was not significantly reduced by citalopram, venlafaxine, or 5-HT at any of the concentrations tested (Fig. 4.5F), suggesting that inhibition of SERT does not diminish canonical Wnt signaling. Overall, the inhibitory impact of the chemical exposures on canonical Wnt signaling was consistent with their adverse effects on EB morphogenesis.

#### **4.3.5 ACTIVATION OF CANONICAL WNT SIGNALING PARTIALLY ALLEVIATES THE ADVERSE EFFECTS OF FLUOXETINE**

We then tested whether the adverse morphogenetic effects of fluoxetine can be rescued by forced activation of canonical Wnt signaling. CHIR99021 is a pharmacological inhibitor of GSK3 and is widely used to activate canonical Wnt signaling in various cell types (Fig. 4.5A). In P19C5 cells, CHIR99021 increased the TOPFLASH signal in a dose-dependent manner in the presence of fluoxetine at the adverse effect concentrations (6 and 10  $\mu$ M), indicating that CHIR99021 can overcome fluoxetine to activate canonical Wnt signaling (Fig. 4.6A). However, high concentrations of CHIR99021 alone also adversely affected EB morphology. EBs treated with CHIR99021 at 1  $\mu$ M or higher were rounder than control EBs, resulting in a significant reduction in EDI (Fig. 4.6B). This finding suggests that excessive and ubiquitous activation of canonical Wnt signaling is also detrimental to proper axial morphogenesis. Therefore, a rescue experiment for fluoxetine needs to be executed using concentrations of CHIR99021 lower than 1  $\mu$ M to avoid overactivation of Wnt signaling.

EBs were treated with fluoxetine (6  $\mu$ M) together with CHIR99021 at 0.1 or 0.3  $\mu$ M, and EB morphology was assessed after four days of culture. Compared to fluoxetine alone, EBs treated with fluoxetine and CHIR99021 were more elongated (Fig. 4.6C). Both 0.1 and 0.3  $\mu$ M of CHIR99021 were able to significantly increase the EDI of fluoxetine-treated EBs, but the 0.1  $\mu$ M treatment appeared to be

more effective (Fig. 4.6D). Although the extent of rescue was partial, this result supports that the inhibition of canonical Wnt signaling is responsible for the adverse morphogenetic effects of fluoxetine.

#### **4.3.6 TRIFLUOROMETHYLPHENYL MOIETY OF FLUOXETINE IS ESSENTIAL IN CAUSING THE ADVERSE EFFECTS OF FLUOXETINE**

XAV939 is a pharmacological inhibitor of tankyrases (TNKS; (Huang *et al.*, 2009), which are negative regulators of AXIN, the rate-limiting component of the  $\beta$ -Catenin destruction complex (Fig. 4.5A). In examining the chemical structure of XAV939 and fluoxetine, we noticed that both molecules contain a trifluoromethylphenyl group (Fig. 4.7A, B). Norfluoxetine, which had similar morphologic effects to its parent molecule, fluoxetine (Fig. 4.2C), also contains a trifluoromethylphenyl group (Fig. 4.7A). Because of the structural similarity, it is possible that fluoxetine may act similarly to XAV939 to inhibit canonical Wnt signaling and alter EB morphogenesis. To assess whether the trifluoromethylphenyl moiety in fluoxetine was necessary in causing its adverse impacts, we examined the effect of nisoxetine, a norepinephrine reuptake inhibitor (NRI). Nisoxetine is structurally similar to fluoxetine, except that a trifluoromethyl group on the aromatic ring (in *para* position) is replaced with a methoxy group (in *ortho* position) (Fig. 4.7A, B). Nisoxetine did not adversely affect EB morphogenesis even at the highest concentration evaluated (20  $\mu$ M; Fig. 4.7C). In addition, nisoxetine did not significantly inhibit canonical Wnt signaling, based on the TOPFLASH assay (Fig. 4.7D). These results suggest the trifluoromethylphenyl moiety in fluoxetine mediates its morphogenetic impact and its inhibitory effect on Wnt signaling.

#### **4.3.7 FLUOXETINE DIMINISHES CELL PROLIFERATION INDEPENDENTLY OF ITS INHIBITORY EFFECTS OF WNT SIGNALING OR SERT**

Despite the similarities in structure and inhibitory effect on canonical Wnt signaling, XAV939 and fluoxetine exhibit different impacts on EB morphogenesis. XAV939 (0.5 - 5  $\mu$ M) causes a marked reduction in EDI without affecting area (Li and Marikawa, 2015), whereas fluoxetine (6 - 10  $\mu$ M) lowered both area and EDI (Fig. 4.2A). A reduction in area indicates that fluoxetine diminishes cell proliferation or viability during EB development. To further assess the cytostatic impact of fluoxetine, P19C5 cells in monolayer were cultured for four days in the presence of fluoxetine at various concentrations, and the number of viable cells was measured. Compared to the control, fluoxetine caused a dose-dependent decrease in the number of viable cells at 6  $\mu$ M and higher (Fig. 4.8A). Note that 6  $\mu$ M of fluoxetine decreased the viable cell number by about 20%, which was comparable to the extent of reduction in EB area at the same concentration of fluoxetine (Fig. 4.2B). Fluoxetine treatment at higher than 10  $\mu$ M, which was detrimental to the survival of EBs, decreased the viable cell number by >95% in monolayer culture. This suggests that cytostatic effects of fluoxetine reduce EB growth.

R-fluoxetine, S-fluoxetine, and norfluoxetine also decreased the number of viable cells in monolayer culture in a dose-response manner comparable to the effect of fluoxetine, although S-fluoxetine again appeared slightly less potent (Fig. 4.8A). In contrast, the viable cell number was not markedly reduced by nisooxetine even at the highest concentration examined (20  $\mu$ M; Fig. 4.8A), suggesting that the trifluoromethylphenyl moiety is essential for the cytostatic effect of fluoxetine. Importantly, XAV939 (up to 2.5  $\mu$ M) did not reduce the number of viable cells, although it markedly suppressed the TOPFLASH signal (Fig. 4.8B). Therefore, neither the inhibition of canonical Wnt signaling nor the possession of a trifluoromethylphenyl group was sufficient to diminish cell proliferation.

Lastly, to test whether the action as an SSRI is linked to the cytostatic effect, the viable cell number in monolayer culture was assessed after exposure to citalopram, venlafaxine, or 5-HT (Fig. 4.8C, D). None of these compounds significantly decreased the viable cell number, suggesting that inhibition of SERT is not responsible for the cytostatic effect of fluoxetine.



## 4.4 - Discussion

In the present study, we used the P19C5 EB morphogenesis model to investigate the developmental toxicity of fluoxetine, a common SSRI antidepressant that is often prescribed to women of reproductive age. EB morphogenesis was adversely affected by fluoxetine and norfluoxetine at 6  $\mu\text{M}$  and above. Comparison to other serotonin reuptake inhibitors suggested that the adverse morphogenetic effects of fluoxetine were not mediated by SERT inhibition. Gene expression analyses in EBs showed that various developmental regulators were affected by fluoxetine, particularly those involved in mesodermal differentiation, and implied inhibition of canonical Wnt signaling. Signaling reporter assays confirmed that fluoxetine inhibits canonical Wnt signaling. Fluoxetine also exhibited cytostatic effects on P19C5 cells independently of its inhibition of SERT and Wnt signaling. We propose that the adverse morphogenetic impacts of fluoxetine reflect its effects on canonical Wnt signaling and cell proliferation rather than its therapeutic effects as a serotonin reuptake inhibitor (Fig. 4.9). Considering the epidemiological findings linking maternal fluoxetine intake with an increased incidence of birth defects, the present study provides mechanistic insight for further investigations into the safety of antidepressant use during pregnancy.

The findings of this study are only relevant if developing embryos *in vivo* are exposed to fluoxetine and norfluoxetine at the levels that cause adverse effects in EBs (i.e., 6  $\mu\text{M}$  or higher). A large multicenter study has shown that the mean plasma concentrations ( $\pm$  standard deviation) of fluoxetine, norfluoxetine, and a sum of the two in depressed patients are  $0.31 \pm 0.16 \mu\text{M}$ ,  $0.41 \pm 0.16 \mu\text{M}$ , and  $0.73 \pm 0.26 \mu\text{M}$ , respectively (Amsterdam *et al.*, 1997). Other studies involving depressed patients or healthy volunteers are also consistent with these plasma concentrations (Brunswick *et al.*, 2002; Cheer and Goa, 2001; Teter *et al.*, 2005). As fluoxetine and norfluoxetine appear to cross the placental barrier readily, *in utero* concentrations may be comparable to the maternal plasma levels (Kim *et al.*, 2004, 2006). If that is the case, exposure levels to developing embryos on average may be one order of magnitude lower than the concentrations that caused adverse morphogenetic effects in the present study. However, fluoxetine and norfluoxetine exhibit tissue-specific accumulation, and their levels in the brain have been measured at concentrations 10 - 20 times higher than the plasma levels (Bolo *et al.*, 2000; Karson *et al.*, 1993). Thus, it is important to determine whether similar accumulation occurs in the embryo or the reproductive tract. Furthermore, other conditions, such as hepatic impairment, intake of other medications that inhibit metabolizing enzymes, and accidental overdose, may significantly increase the plasma concentrations of fluoxetine and norfluoxetine (Cheer and Goa, 2001). Therefore, although the average fluoxetine concentration during pregnancy may be below the adverse effect levels, many factors may temporarily increase exposure levels in certain individuals and contribute to the higher incidence of congenital malformations.

Fluoxetine inhibited canonical Wnt signaling in P19C5 cells. Although the exact mechanism by which fluoxetine interferes with Wnt signaling is unclear, the structural similarity to XAV939 raises the possibility that fluoxetine may act as an inhibitor of TNKS (tankyrase) to stabilize AXIN (Fig. 4.5A). Inhibition of canonical Wnt signaling by fluoxetine was also reported in a recent study using chondrogenic cells, in which the TOPFLASH signal and Wnt-mediated osteoarthritis-related phenotypes are attenuated by fluoxetine treatment (Miyamoto *et al.*, 2017). The concentrations of fluoxetine that inhibited Wnt signaling in these cells (e.g.,  $\geq 5 \mu\text{M}$  to reduce the TOPFLASH signal) are comparable to the fluoxetine exposures ( $\geq 6 \mu\text{M}$ ) that diminished Wnt signaling activity and altered EB morphology in the present study. Miyamoto *et al.* (2017) also demonstrated enhanced binding of  $\beta$ -Catenin (CTNNB1) to AXIN in cell lysate in the presence of fluoxetine, suggesting that fluoxetine may directly inhibit canonical Wnt signaling by stabilizing the association of  $\beta$ -Catenin with the destruction complex. Conversely, another study using hippocampal neural precursor cells (NPC) has reported that fluoxetine activates, rather than inhibits, canonical Wnt signaling, as evidenced by increases in nuclear  $\beta$ -Catenin amount and the TOPFLASH signal (Hui *et al.*, 2015). Note, however, that the activation of canonical Wnt signaling in NPCs is induced by fluoxetine treatment at  $1 \mu\text{M}$ , which is substantially lower than the levels that inhibited Wnt signaling in chondrogenic and P19C5 cells. Interestingly, the accumulation of nuclear  $\beta$ -Catenin and Wnt-induced cell proliferation are attenuated in NPCs by co-treatment with WAY-100635, an antagonist of the 5-HT receptor, which implies that the activation of Wnt signaling is caused by the SSRI action of fluoxetine (Hui *et al.*, 2015). Thus, the apparent contradiction on how fluoxetine modulates Wnt signaling might be explained by the ability of NPCs to respond to the elevated 5-HT level caused by inhibition of SERT, which appears to be absent in P19C5 cells. If fluoxetine can directly act on the signaling components, such as TNKS, AXIN and  $\beta$ -Catenin, then higher concentrations of fluoxetine might inhibit canonical Wnt signaling in NPCs as well.

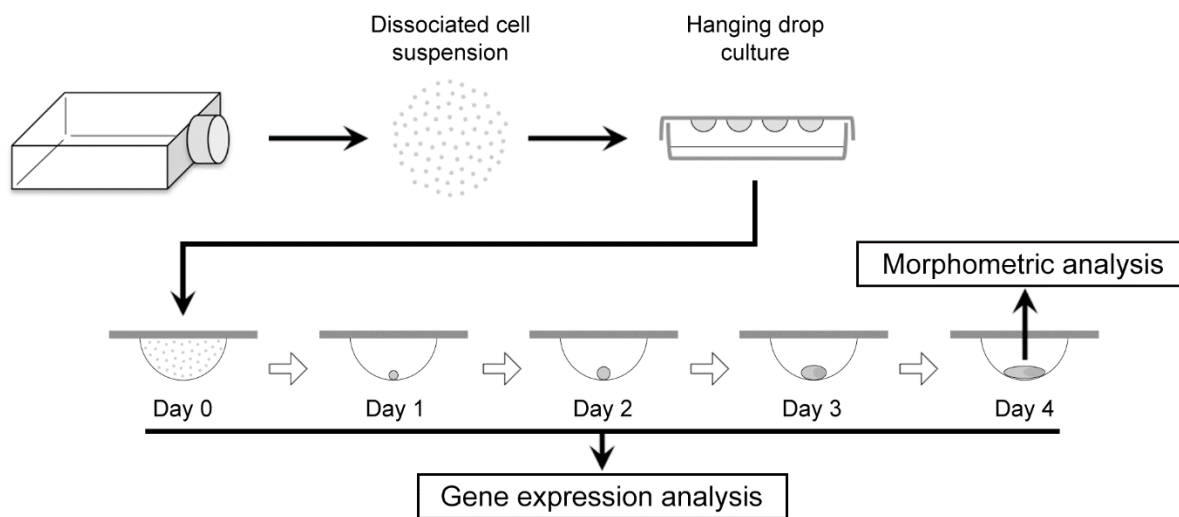
Inhibition of canonical Wnt signaling may contribute to the etiology of birth defects associated with maternal fluoxetine intake, specifically cardiac malformations, such as ventricular septal defect (VSD), atrioventricular septal defect (AVSD), and right ventricular outflow tract obstruction (RVOTO) (Elfolk and Malm, 2010; Gao *et al.*, 2017; Kioussi *et al.*, 2002; Reefhuis *et al.*, 2015; Ruiz-Villalba *et al.*, 2016; Grigoryan *et al.*, 2008; Azhar and Ware, 2016). Genetic studies in mouse embryos have demonstrated that canonical Wnt signaling plays crucial roles in the development of the heart (Gessert and Kühl, 2010; Ruiz-Villalba *et al.*, 2016; Grigoryan *et al.*, 2008; Azhar and Ware, 2016). Functional  $\beta$ -Catenin is required for the differentiation of cardiac progenitors (Lin *et al.*, 2007), the growth and diversification of the second heart field precursors into right ventricular and interventricular myocardium (Ai *et al.*, 2007; Kwon *et al.*, 2007; Klaus *et al.*, 2007), the migration of cardiac neural crest (Kioussi *et al.*, 2002), and the formation of the septum from the endocardial cushion (Liebner *et al.*, 2004). Misregulation of these events during heart development may result in cardiac malformations, such as VSD, AVSD, and RVOTO (Lalani and Belmont, 2014). Note, however, that these genetic studies are mainly based on the loss of  $\beta$ -Catenin in

specific embryonic tissues, which probably abolishes the activity of canonical Wnt signaling completely. In contrast, Wnt signaling inhibition by fluoxetine at 6-10  $\mu$ M appeared to be more modest than the loss of  $\beta$ -Catenin, judging from the extent of the TOPFLASH signal reduction (Fig. 4.5B). Further investigations are warranted to understand how moderate inhibition of canonical Wnt signaling affects the developing heart.

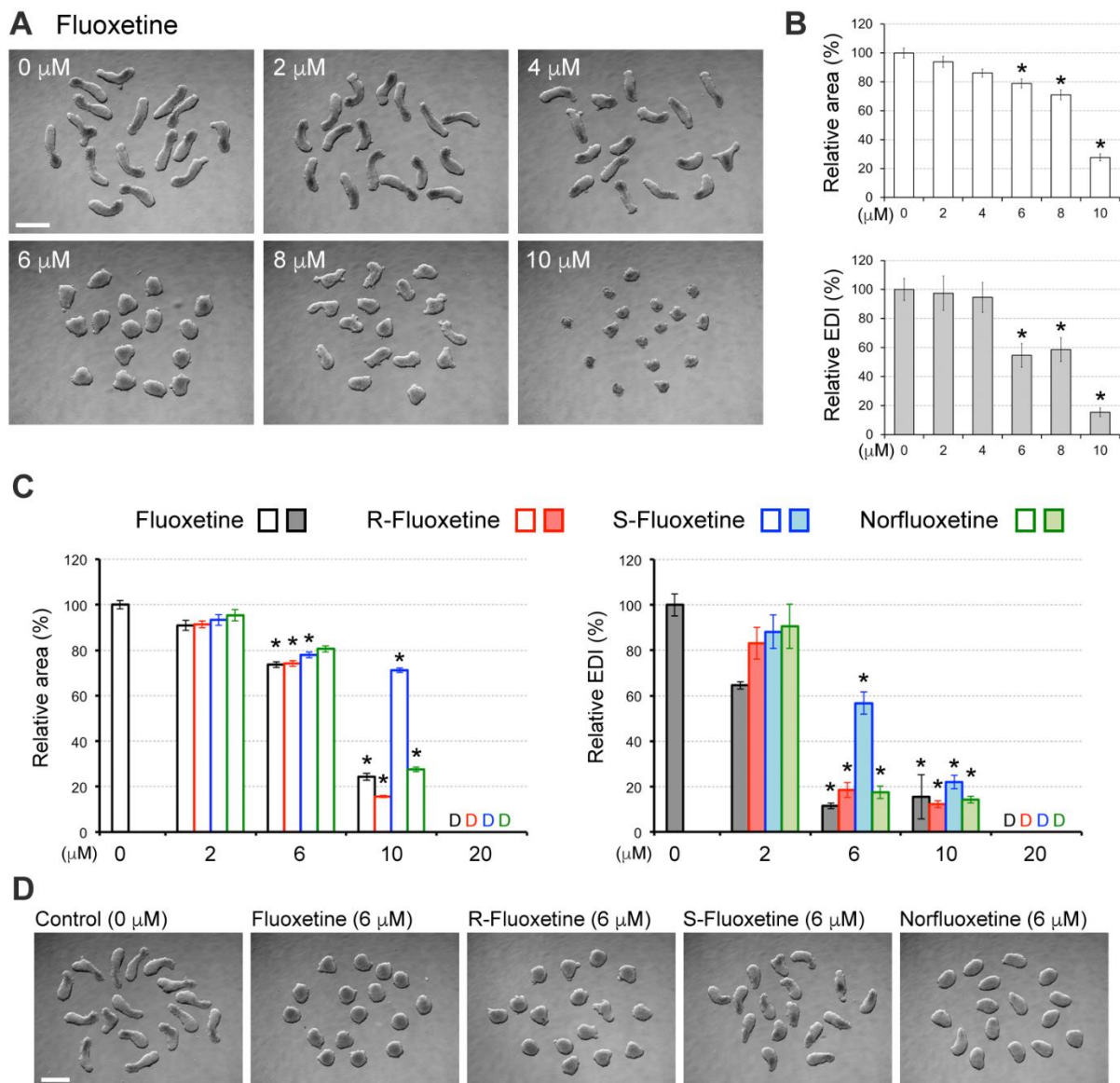
The reduction of EB growth caused by fluoxetine was likely due to its cytostatic activity independently of its inhibitory effects on SERT and canonical Wnt signaling. Fluoxetine diminishes cell proliferation and viability in a variety of cell types, including breast cancer cell lines (Bowie, 2015; Sun *et al.*, 2015), ovarian cancer cell lines (Lee *et al.*, 2010), neuroblastoma cell lines (Choi *et al.*, 2017), pheochromocytoma cells (Han and Lee, 2009), hippocampal cells (Post *et al.*, 2000; Schaz *et al.*, 2011), and hypothalamic neuroprogenitor cells (Sousa-Ferreira *et al.*, 2014), at concentrations that correspond to the adverse effect levels in P19C5 cells. The negative impact of fluoxetine on cellular proliferation and/or survival has been hypothesized to occur via a number of molecular pathways, including induction of endoplasmic reticulum (ER) stress, generation of reactive oxygen species (ROS), activation of NF- $\kappa$ B, increased influx of extracellular calcium, and phosphorylation of mitogen-activated kinases (MAPKs) (Lee *et al.*, 2010; Post *et al.*, 2000; Choi *et al.*, 2017). The anti-proliferative effects of fluoxetine may not be mediated by inhibition of SERT in some cases. In the pheochromocytoma cell line, fluoxetine enhances the effect of a cytotoxic agent (1-methyl-4-phenylpyridinium), whereas another SERT inhibitor amitriptyline attenuates the cytotoxic effect (Han and Lee, 2009). In the human embryonic kidney cell line, the dose-response profiles of the anti-proliferative effects of fluoxetine are similar between unmanipulated cells (no SERT expression) and those with SERT overexpression (Schaz *et al.*, 2011). The mechanism by which fluoxetine interacts with the molecular pathways described above in an SERT-independent manner is still unclear. Nonetheless, it is of interest to investigate the involvement of these pathways in the fluoxetine-induced reduction in P9C5 EB growth and the inhibition of canonical Wnt signaling.

In addition to fluoxetine, various other types of SSRIs are frequently prescribed to reproductive-age women. Epidemiologic studies that compiled data for all SSRIs together as a single class indicate the absence of a significant association between maternal SSRI intake and the occurrence of birth defects (Ornoy and Koren, 2017). However, fluoxetine may cause birth defects through mechanisms unrelated to its SSRI action, as suggested in the present study. For this reason, it is crucial to investigate the developmental toxicity of individual medications with consideration for unexpected off-target effects. Currently, animal-based tests are the gold standard to assess the developmental toxicity of chemical compounds. However, animal experimentation is costly, labor-intensive, and it poses ethical issues concerning animal welfare, especially given the large number of compounds that need to be tested. *In vitro* assays, such as the P19C5 EB morphogenesis-based tests, may be effectively used to gain valuable information on dose-response relationships and molecular mechanisms for a number of compounds in a practical and ethical manner. Our ongoing investigation using P19C5 EBs suggests that a few other

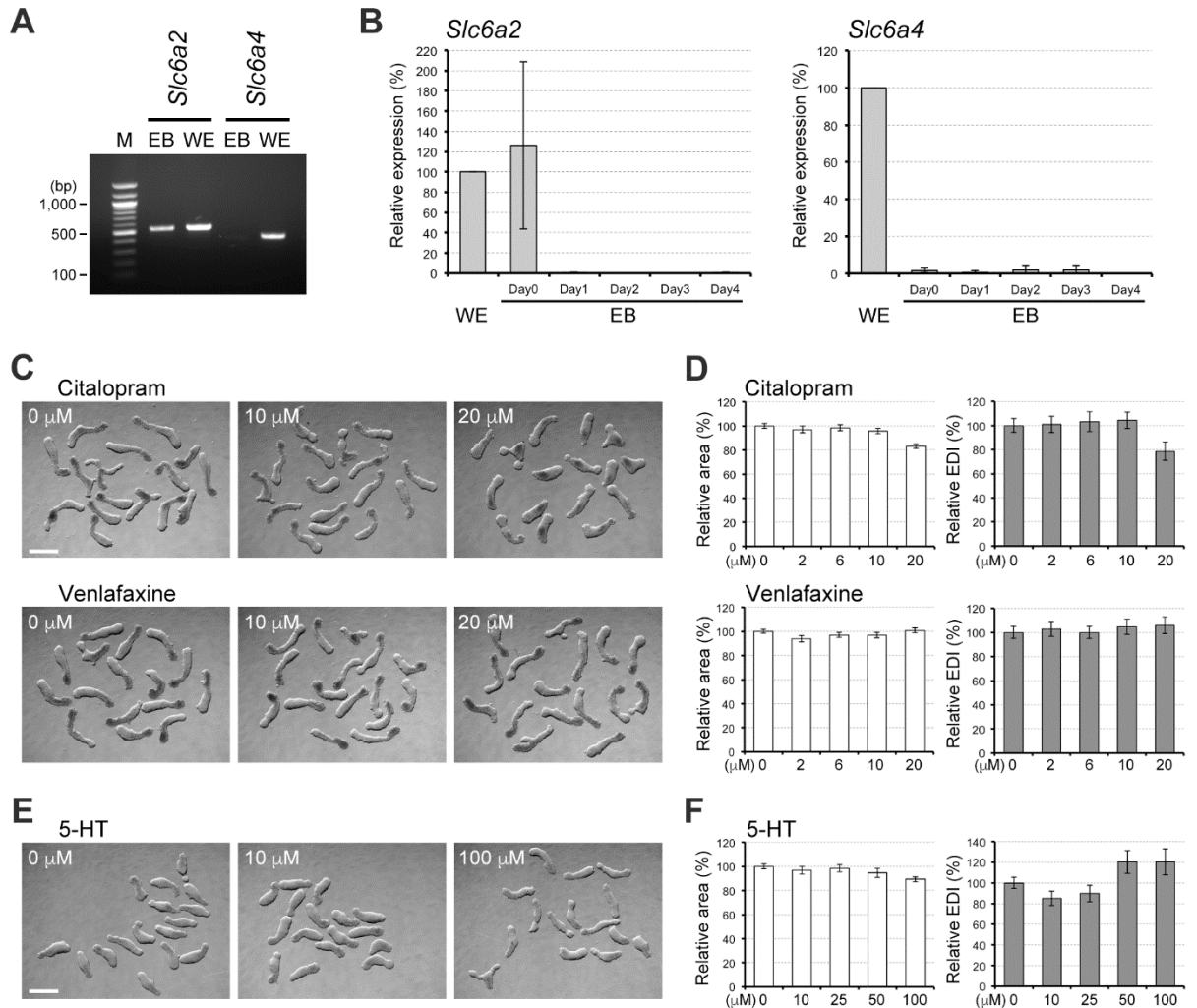
commonly prescribed antidepressants, notably paroxetine and sertraline, also have adverse effects on developmental processes. Tests like this are important because, with a better understanding of the properties of individual medications, physicians can choose the most appropriate options to treat depressed women while minimizing the risk of causing unnecessary birth defects in human babies.



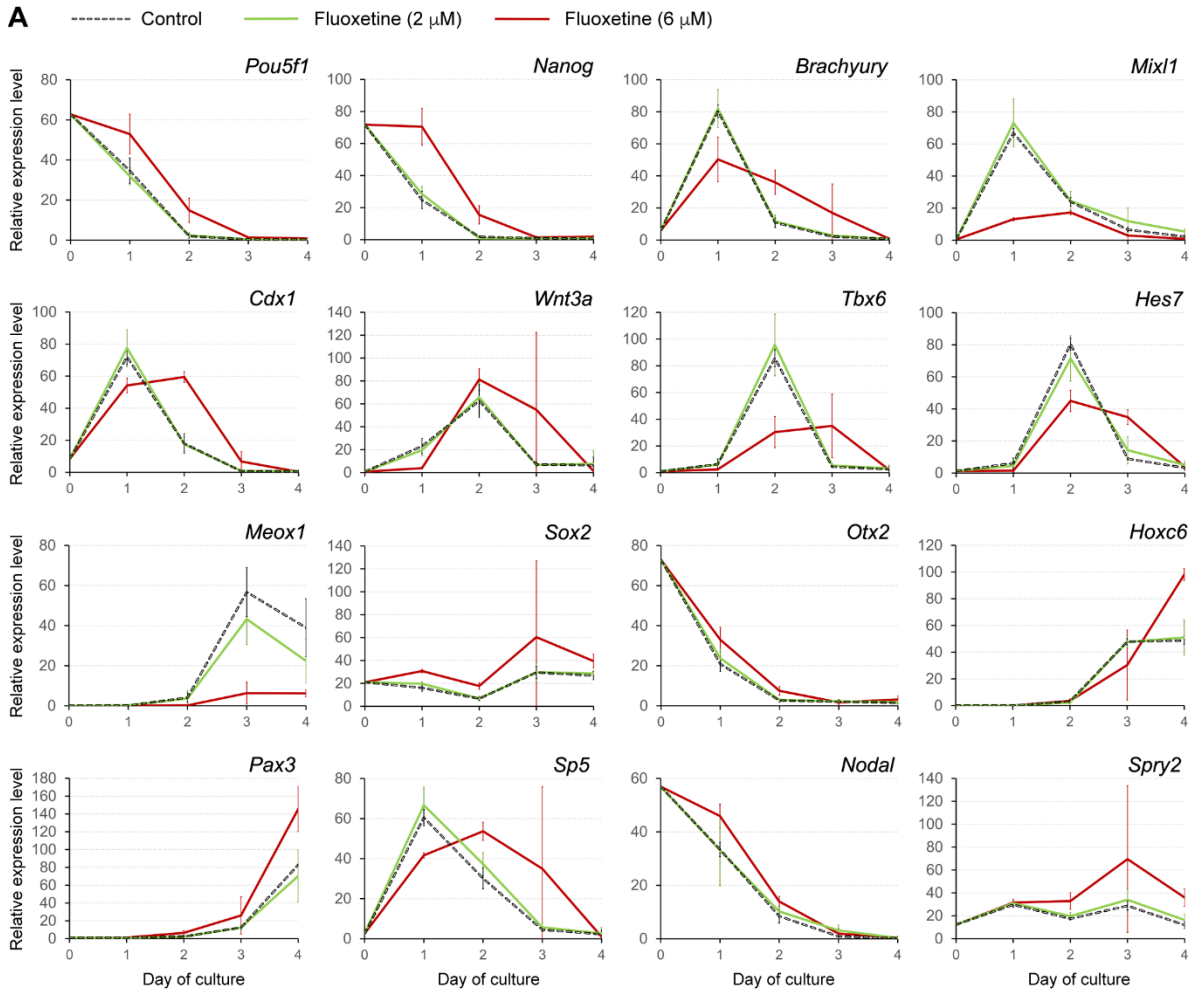
**Figure 4.1.** The experimental scheme to assess the morphogenetic and molecular impact of chemical exposures using the P19C5 embryoid body morphogenesis system.



**Figure 4.2.** Fluoxetine impairs morphogenesis of P19C5 embryoid bodies. (A) Representative images of embryoid bodies (EBs) on Day 4 that have been treated with fluoxetine at the indicated concentrations. All images show a set of EBs that were generated from the same cell suspension. (B) Morphometric parameters of Day 4 EBs treated with fluoxetine. Graphs show averages of relative area (open columns) and relative EDI (shaded columns) with error bars representing the 95% confidence intervals. Asterisks indicate adverse impacts, which are defined as a reduction in average relative area by > 20% and/or a change in average relative EDI by > 40%, compared to the control (0 μM). All adverse impacts are statistically significant ( $P < 0.01$ ; two-sample t-test). (C) Morphometric parameters of Day 4 EBs treated with fluoxetine's R- or S- enantiomers or primary metabolite, norfluoxetine, presented in the same format as described for (B). No area or EDI values are available when EBs are dead (noted as "D"). (D) Representative images of control and compound-treated Day 4 EBs. Scale bars in (A, D) = 500 μm.



**Figure 4.3.** The morphogenetic effects of fluoxetine are not due to inhibition of the serotonin transporter. (A) RT-PCR products of the *Slc6a2* (norepinephrine transporter) and *Slc6a4* (serotonin transporter) mRNA, amplified from the embryoid body (EB; a mixture of Days 0, 1, 2, 3, and 4) and the mouse whole embryo (WE; embryonic stage E10.5) cDNAs. M: 100 bp ladder molecular marker. (B) Quantitative RT-PCR analysis of *Slc6a2* and *Slc6a4* expressions in EBs, relative to the level in WE. Graphs show averages of relative expression levels, and error bars show standard deviation. (C) Representative images of Day 4 EBs treated with citalopram or venlafaxine. (D) Morphometric parameters of Day 4 EBs treated with citalopram or venlafaxine. (E) Representative images of Day 4 EBs that were treated with serotonin (5-HT). Morphometric parameters of Day 4 EBs treated with 5-HT. Graphs in (D, F) show averages of relative area and relative EDI with error bars of 95% confidence intervals. Scale bars in (C, E) = 500  $\mu$ m.

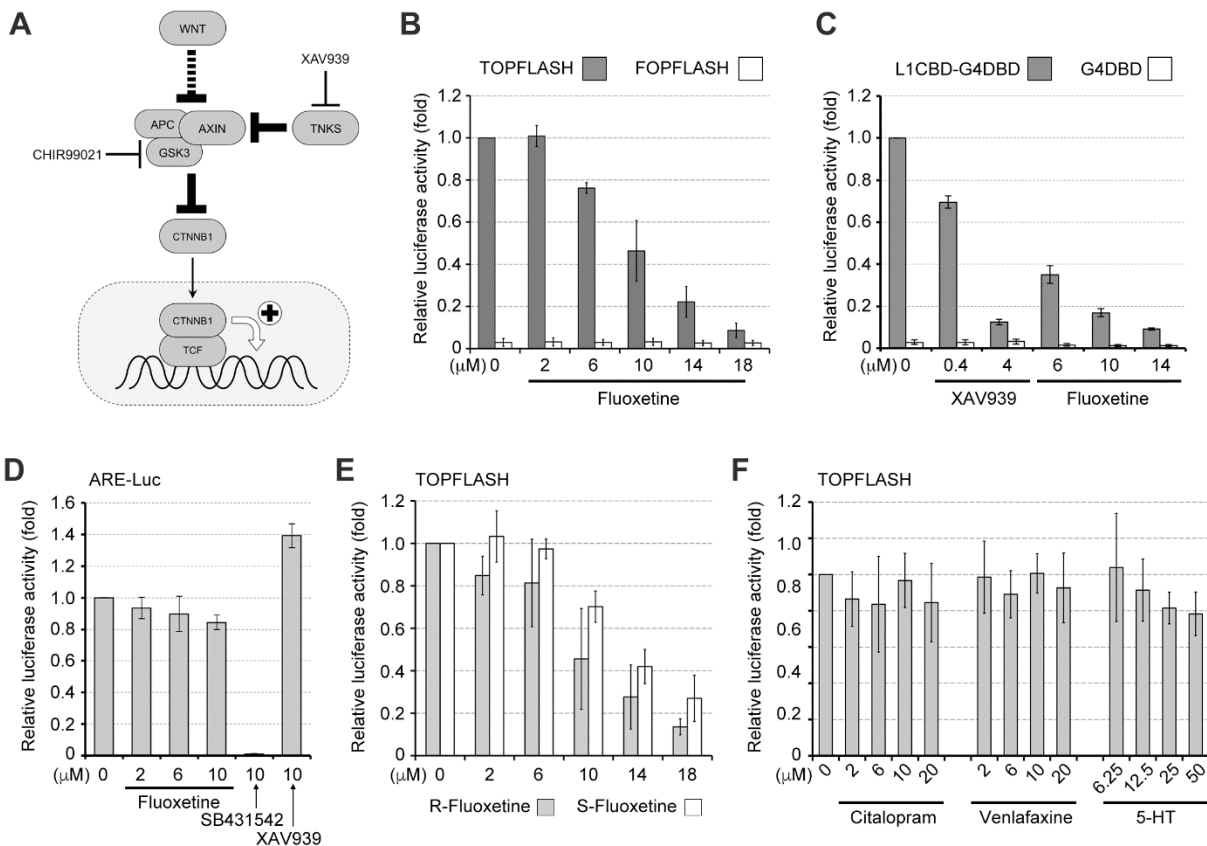


**B**

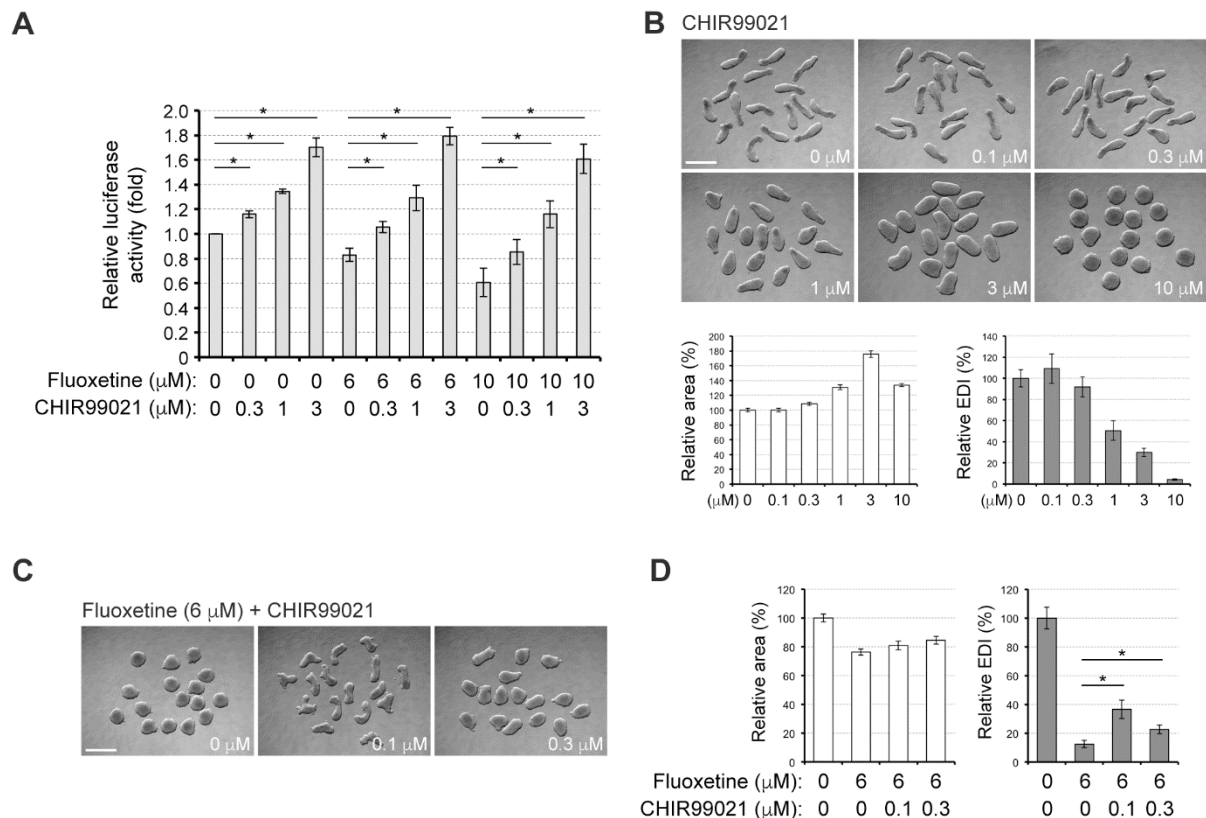
	Inhibited signal	<i>Pou5f1</i>		<i>Nanog</i>		<i>Brachyury</i>	
		Day 1	Day 2	Day 1	Day 2	Day 1	Day 2
XAV939 (5 $\mu$ M)	WNT	▲	▲	▲	▲	▼	▲
SB431542 (10 $\mu$ M)	NODAL	NS	NS	NS	NS	▼	▼
PD173074 (100 nM)	FGF	NS	NS	▲	NS	▼	NS
DMH1 (200 nM)	BMP	NS	NS	NS	NS	NS	NS
BMS493 (1 $\mu$ M)	RA	NS	NS	NS	NS	NS	NS
Fluoxetine (6 $\mu$ M)		▲	▲	▲	▲	▼	▲

**Figure 4.4.** Fluoxetine alters the expression patterns of various developmental regulator genes. (A) Temporal expression profiles of developmental regulator genes during EB development, determined by quantitative RT-PCR analysis. EBs were treated with fluoxetine (2 or 6  $\mu$ M) or DMSO (control). Horizontal axes represent days of culture, and vertical axes represent relative expression levels in arbitrary units. Graphs show average relative expression levels, and error bars represent standard deviation. (B) Comparisons of expression profiles for *Pou5f1*, *Nanog*, and *Brachyury* on Days 1 and 2 between EBs treated with fluoxetine and EBs treated with pharmacological inhibitors of the major developmental signals. The disruption profiles for the WNT, NODAL, FGF, BMP, and retinoic acid (RA) signaling inhibitors are based on the previous study (Li and Marikawa, 2015). NS: No significant change.

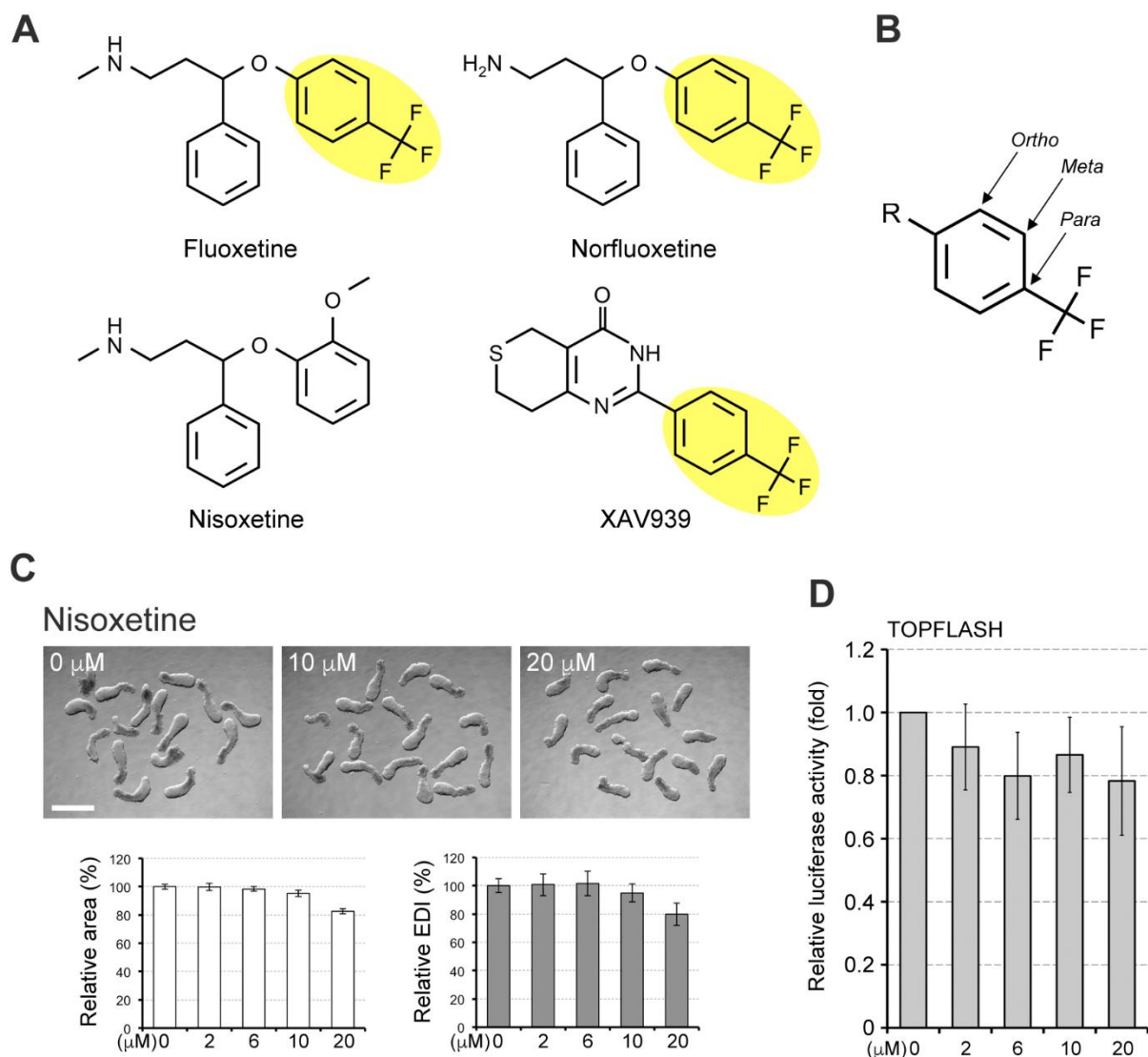




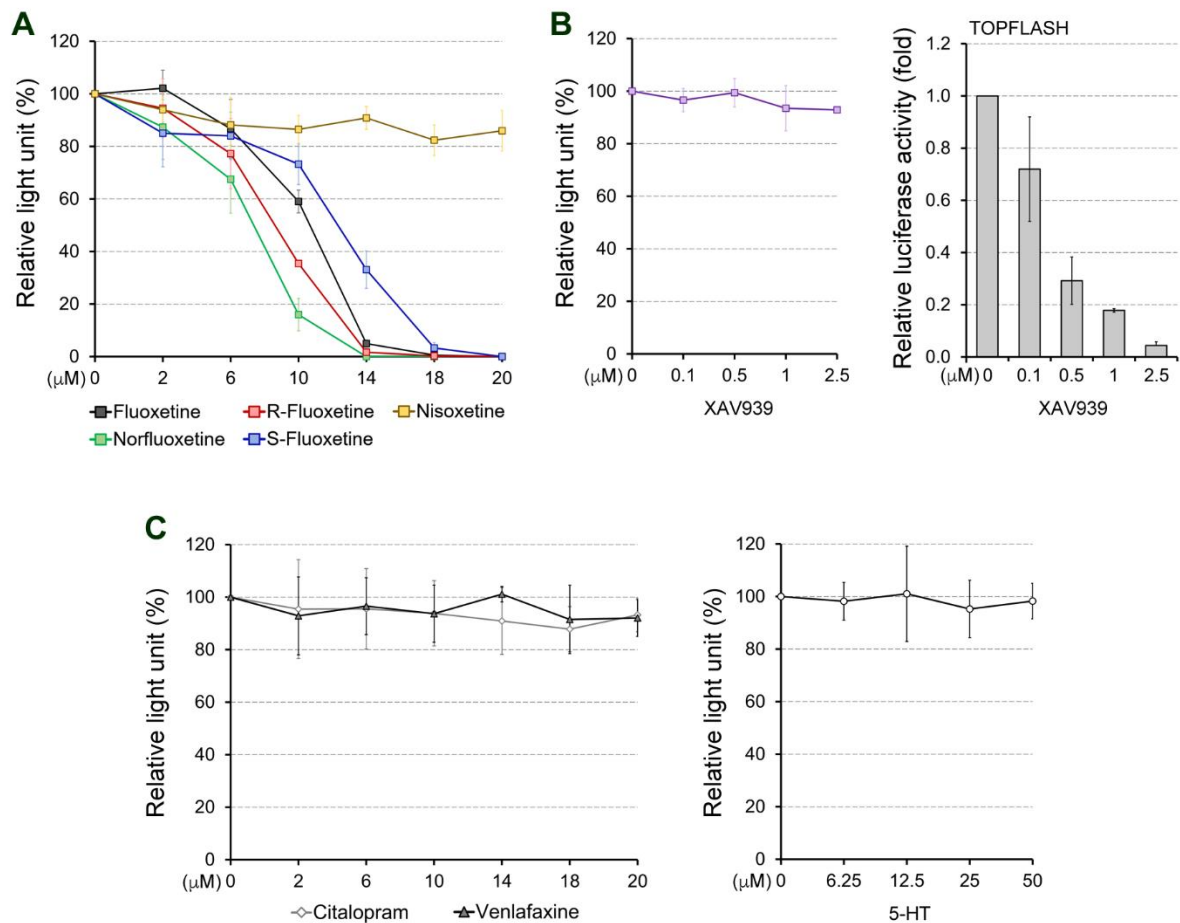
**Figure 4.5.** Fluoxetine inhibits canonical Wnt signaling in P19C5 cells. (A) A schematic diagram of the canonical Wnt signaling pathway, highlighting the key regulatory components and pharmacological inhibitors (XAV939 and CHIR99021). (B, C, D) Relative luciferase activity of TOPFLASH and FOPFLASH (B), pG5-Luc driven by L1CBD-G4DBD and G4DBD (C), and ARE-Luc (D), in P19C5 cells in monolayer culture, treated with the indicated compounds for 24 hours. (E, F) Impact of fluoxetine R- or S-enantiomers (E), citalopram, venlafaxine, and 5-HT (F) on the relative luciferase activity of TOPFLASH in monolayer culture. All graphs show averages of relative luciferase activities with error bars of standard deviation.



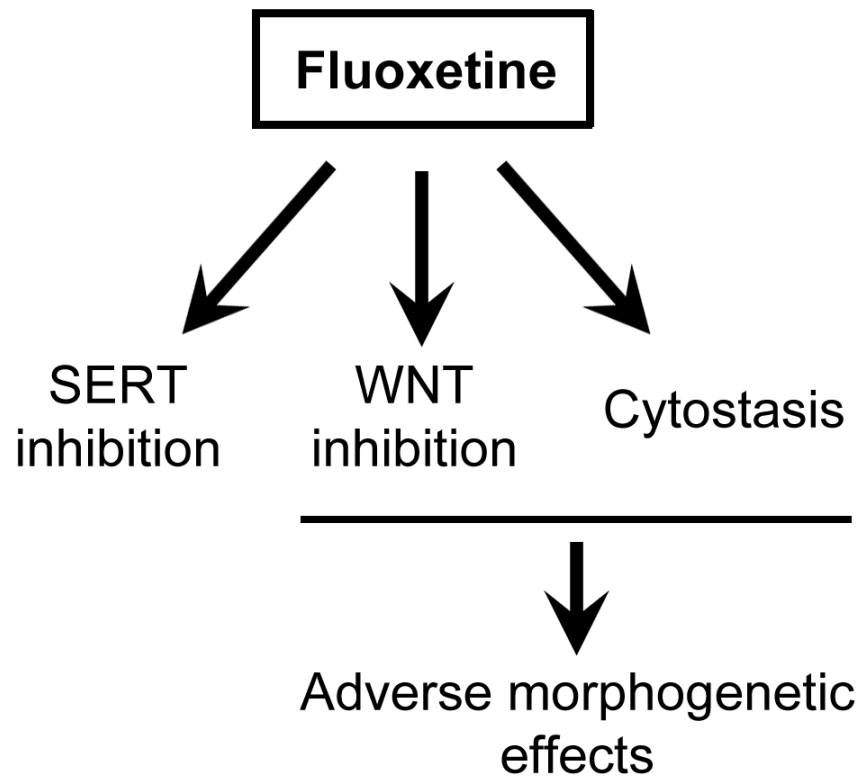
**Figure 4.6.** Activation of canonical Wnt signaling partially alleviates the adverse effects of fluoxetine. (A) Relative luciferase activity of TOPFLASH in monolayer P19C5 cells treated with fluoxetine and CHIR99021 at various concentrations for 24 hours. (B) Representative images (top) and morphometric parameters (bottom) of Day 4 EBs treated with CHIR99021. Graphs show the averages of relative area and relative EDI and error bars represent the 95% confidence intervals. (C) Representative images of Day 4 EBs treated with fluoxetine and CHIR99021. (D) Morphometric parameters of Day 4 EBs treated with fluoxetine and CHIR99021, presented in the same format as described for (B). Asterisks in (A, D) indicate statistically significant differences between the two groups indicated ( $P < 0.01$ ; two-sample t-test). Scale bars in (B,C) = 500  $\mu\text{m}$ .



**Figure 4.7.** Trifluoromethylphenyl moiety of fluoxetine is essential in causing the adverse effects of fluoxetine. (A) Chemical structures of fluoxetine, norfluoxetine, nisoxetine, and XAV939. Trifluoromethylphenyl groups are highlighted in yellow. (B) Structure of the trifluoromethylphenyl group, in which a trifluoromethyl substituent is on the aromatic ring in the *para* position relative to R. (C) Representative images (top) and morphometric parameters (bottom) of Day 4 EB treated with nisoxetine. Graphs show averages of relative area and relative EDI with error bars showing 95% confidence intervals. Scale bar = 500  $\mu\text{m}$ . (D) Relative luciferase activity of TOPFLASH in P19C5 cells treated with nisoxetine for 24 hours in monolayer culture. Graphs show average  $\pm$  standard deviation of relative luciferase activity.



**Figure 4.8.** Fluoxetine diminishes cell proliferation independently of its inhibitory effects on Wnt signaling or SERT activity. (A) Impact of fluoxetine, R-fluoxetine, S-fluoxetine, norfluoxetine, and nisoxetine on the number of viable cells in monolayer culture after 4 days of treatment, assessed by the CellTiter-Glo Luminescent Assay. (B) Impact of XAV939 on the number of viable cells (left; 4 day treatment) and on the TOPFLASH activity (right; 24 hour treatment) of P19C5 cells in monolayer culture. (C) Impact of citalopram and venlafaxine (left), and 5-HT (right), on the number of viable cells in monolayer culture after 4 days of treatment. All graphs show average and standard deviation of relative light unit or luciferase activity.



**Figure 4.9.** A proposed mechanism of action by which fluoxetine adversely affects morphogenesis.

**Table 4.1.** Genes examined in the fluoxetine study.

Gene name	Characteristics *1	Expression in EBs *2	Primer sequences (5' → 3')
<i>Actb</i>	Cytoplasmic actin; Constitutively expressed house keeping gene; Used as a loading control to normalize the expression levels of other genes	Constitutive	F: GAGAGGGAAATCGTGCGTGACATC R: CAGCTCAGTAACAGTCCGCCTAGA
<i>Brachyury</i>	T-box transcription factor; Associated with the initiation of gastrulation; Expressed in the early primitive streak	Day 1 peak	F: CCTCGGATTACATCGTGAGAGTT R: AGTAGGTGGGCGGGCGTTATGACT
<i>Cdx1</i>	Homeodomain transcription factor; Expressed in the early primitive streak; Regulator of cranial-caudal axial patterning	Day 1 peak	F: TCAGGACTGGACATGAGGTAGAGG R: TGGAAGGTGGGCATGAGCAGGTA
<i>Hes7</i>	BHLH transcription factor; Expressed in presomitic mesoderm; Transcriptional target of Notch signaling	Day 2 peak	F: CATACCCTTCTCCACCTTTAGGC R: AGTGACGAGAAAGCGATTCAAAGG
<i>Hoxc6</i>	Homeodomain transcription factor; Regulator of cranial-caudal axial patterning; Expressed in the neural tube	Up-regulation toward Day 4	F: TTCGCCACAGGAGAATGTCGTGTT R: CGAGTTAGGTAGCGGTTGAAGTGA
<i>Meox1</i>	Homeodomain transcription factor; Regulator of somite segmentation; Expressed in somitic mesoderm	Day 3 peak	F: AAAATCAGACTTCCACGCGACAG R: TTCACACGTTTCCACTTCATCCTC
<i>Mixl1</i>	Homeodomain transcription factor; Associated with the initiation of gastrulation; Expressed in the early primitive streak	Day 1 peak	F: CGACAGACCATGTACCCAGACATC R: TGAGGCTTCAAACACCTAGCTTCA
<i>Nanog</i>	Homeodomain transcription factor; Regulator of pluripotency maintenance; Expressed in the epiblast	Down-regulation after Day 0	F: GCTTTGGAGACAGTGAGGTGCATA R: GCTACCTCAAACCTCTGGTCCTT
<i>Nodal</i>	Ligand for Nodal signaling; Transcriptional target of Nodal signaling	Down-regulation after Day 0 or 1	F: GTACATGTTGAGCCTCTACCGAGA R: TCTACAGACAGCTGTCCCTCCTGG
<i>Otx2</i>	Homeodomain transcription factor; Expressed in the neural tube	Down-regulation after Day 0 or 1	F: GAAACAGCGAAGGGAGAGGACGAC R: CCCAAAGTAGGAAGTTGAGCCAGC
<i>Pax3</i>	Paired-box transcription factor; Expressed in the neural tube	Up-regulation toward Day 4	F: GCTTCTCAGCGTGCAATACTGTGT R: TTTCTGTTCTAGCCCTGCCTTTTG
<i>Pou5f1</i>	POU-domain transcription factor; Regulator of pluripotency maintenance; Expressed in the epiblast	Down-regulation after Day 0	F: AGGCAGGAGCACGAGTGGAAGCA R: GGAGGGCTTCGGGCACTTCAGAAA
<i>Sox2</i>	SRY-box transcription factor; Regulator of pluripotency maintenance; Expressed in the epiblast and neural tube	Down-regulation at Day 2	F: CACATGAAGGAGCACCCGGATTAT R: CTGGAGTGGGAGGAAGAGGTAACC
<i>Sp5</i>	Zinc finger transcription factor; Transcriptional target of canonical Wnt signaling; Expressed in the early primitive streak	Day 1 peak	F: CAGGACAGGAAACTGGGTGCTAGT R: GGCCTAGCAAAAACCTTAGGCCTTG
<i>Spry2</i>	Antagonist of RTK signaling; Transcriptional target of FGF signaling	Up-regulation after Day 0	F: TTGCATAAGCTAAAGCAGCCAACA R: TTTGTGACTGTGCCATGAAGCATA
<i>Tbx6</i>	T-box transcription factor; Regulator of paraxial mesoderm specification; Expressed in the caudal end	Day 2 peak	F: GGCCTCTCTCCACCCTTTAGTTC R: CACTAGTAACAAGGCCCCCAGGAG
<i>Wnt3a</i>	Ligand of the Wnt signaling; Regulator of paraxial mesoderm specification; Expressed in the caudal end	Day 2 peak	F: GCCACAAGAGCTTCTGATTGGTA R: CCAGGCAGAAGACAGTCAGTCACC

\*1. Functional and structural properties of the encoded proteins and the expression patterns in normal mouse embryos around the time of gastrulation;

\*2. Temporal expression profiles in unmanipulated (control) P19C5 EBs based on the previous studies (Lau and Marikawa, 2014; Li and Marikawa, 2015, 2016) and the present study.

**Table 4.2.** Compounds used in the fluoxetine study.

<b>Compound name</b>	<b>CASRN</b>	<b>Vendor (catalog number)</b>	<b>Stock concentration (solvent)</b>
Fluoxetine hydrochloride	56296-78-7	Sigma-Aldrich (F132)	50 mM (DMSO)
R-(–)-Fluoxetine hydrochloride	114247-09-5	Sigma-Aldrich (F1678)	20 mM (DMSO)
S-(+)-Fluoxetine hydrochloride	114247-06-2	Sigma-Aldrich (F1553)	20 mM (DMSO)
Norfluoxetine hydrochloride	57226-68-3	Sigma-Aldrich (F133)	10 mM (DMSO)
Citalopram hydrobromide	59729-32-7	Sigma-Aldrich (C7861)	50 mM (DMSO)
Venlafaxine hydrochloride	99300-78-4	Sigma-Aldrich (V7264)	50 mM (H <sub>2</sub> O)
5-Hydroxytryptamine hydrochloride	153-98-0	Sigma-Aldrich (H9523)	50 mM (DMSO)
XAV939	284028-89-3	Sigma-Aldrich (X3004)	10 mM (DMSO)
SB431542	49843-98-3	StemCell Technologies (72232)	10 mM (DMSO)
CHIR99021	252917-06-9	Stemgent (04-0004)	10 mM (DMSO)

CASRN, Chemical Abstracts Service Registry Number; DMSO, dimethyl sulfoxide;

## CHAPTER 5. CONCLUSION

### 5.1 - The future of developmental toxicology research is *in vitro*

The flaws of the current animal-based developmental toxicity screens necessitate that *in vivo* testing be augmented or supplanted by *in vitro* systems, which are fast, cheap and robust to a variety of experimental techniques. The cost of *in vivo* testing alone is a significant factor making *in vitro* methods desirable, because the reliance on animal-based testing makes drug development prohibitively expensive for all but the largest pharmaceutical companies, resulting in a regulatory bottleneck that prevents potentially life-saving therapies from reaching patients and emphasizing the profitability of pharmaceuticals over medical need. To address a small part of this issue, the FDA relaxed the requirement for regulatory testing for “orphan drugs,” medications that are developed to treat, diagnose or prevent rare illnesses. Globally, 6-8% of the population is affected by one of approximately 7,000 rare diseases. The prohibitive cost of drug development means that very few treatments for these diseases are brought to market (Gammie *et al.*, 2015; Blankart *et al.*, 2011; Franco, 2013). Although reducing the regulatory requirements for this small group of drugs may help a few beneficial therapies reach patient populations, the vast majority of potential pharmaceuticals face a gauntlet of regulatory testing. To reduce the cost of drug development and promote the search for new beneficial medications, we should emphasize increasing the ease and efficiency of regulatory test methods using the myriad *in vitro* technologies that have been developed since 1966 rather than depending solely on outdated animal-based methods. Additionally, the relevance of *in vivo* DART testing results to human pregnancies is not always clear. As seen with the thalidomide tragedy, the “apical endpoints” that result from *in vivo* DART testing, such as fetal death or malformation, are not well-suited for comparisons between species and sometimes are inconsistent even within strains of the same species (Tonk *et al.*, 2015). However, since the mechanisms that regulate embryogenesis—particularly within mammals—are highly conserved, “*it is anticipated that mechanistic information on the interference of chemicals with embryogenesis on the molecular level would provide a more informative background for hazard and risk assessment for man...such approaches not only allow a more detailed insight into mechanisms of dysmorphogenesis in animals, but also facilitate direct comparison with the human situation*” (Tonk *et al.*, 2015). Although it is difficult to elucidate teratogenic mechanisms in animal-based models, *in vitro* tests are well suited for this type of hypothesis-based developmental toxicity research.

*In vitro* tests are advantageous in many ways. Relative to *in vivo* testing, *in vitro* test methods make it easier to standardize experimental conditions and allow more objective quantification of results. For DART testing, pregnant animals are usually dosed by body weight, which, combined with individual differences in maternal metabolism and discrepancies in experimenter techniques, leads to large variability in the actual fetal exposure to the treatment chemical. *In vitro* systems allow better control of experimental variables and more accurate treatment doses. Similarly, the outcomes of DART studies are



obtained by semi-quantitative scoring of fetal malformations, which requires extensive experimenter training and potential for human error. *In vitro* systems allow objective quantification of experimental endpoints and reduce inter-laboratory variability. Additionally, while *in vivo* methods are cumbersome and low-throughput, *in vitro* models allow high-throughput screening of hundreds or thousands of chemicals, a scalability that is increasingly needed for regulatory testing. Finally, *in vitro* systems are advantageous because they provide a platform for many different types of experimentation that are difficult or impossible using animal models. Using multiple stem cells lines, it is easy to assess whether a chemical is generally toxic or whether it causes toxicity in a specific type of cell. Additionally, embryonic stem cells can be directed to differentiate into specific tissue types or recreate certain processes of embryonic development *in vitro*. This allows a better understanding of teratogenic susceptibility to a given chemical and makes it easier to extrapolate to humans.

Perhaps the most useful aspect of *in vitro* developmental toxicity testing, particularly in mechanistic evaluations, is the use of gene expression analyses. In contrast to the high volume of chemicals tested using high-throughput methods, gene expression analyses, a.k.a., toxicogenomics, enable high-content screening of potentially teratogenic chemicals. This is a major component of the SEURAT ('Safety Evaluation Ultimately Replacing Animal Testing') initiative and was discussed extensively in the publication that summarized results of the SEURAT first phase: "*Methods such as toxicogenomics are attractive because virtually all toxic responses are preceded by changes in gene expression, and the pattern of gene expression can be diagnostic of mode of action and 'critical biological targets'. Furthermore, because gene expression analysis usually covers the entire genome, the methodology is able to detect off-target effects provided that the offending target or pathway is expressed in the model system(s) being tested. There is already a reasonable amount of research to show that toxicogenomic responses from in vitro systems are rich enough to recapitulate in vivo responses, but more needs to be done to fully define the potential and limitations of in vitro systems as the platforms for toxicogenomics*" (Daston *et al.*, 2015). High-content assays allow investigation of biological responses to developmental toxicants at cellular and molecular levels and are increasingly valuable as a method of understanding teratogenic mode of actions (Knudsen *et al.*, 2011). Between the breadth of analysis offered by high-throughput methods and the depth of mechanistic understanding achieved with high-content toxicogenomic assays, *in vitro* models are changing the field of developmental toxicology for the better.

The dose of chemical used in developmental toxicity testing remains an important consideration that has been difficult to address with *in vivo* DART methods. Teratogens are not always teratogenic. The developmental toxicity of a chemical varies based on the magnitude and timing of embryonic exposure. Ideally, developmental toxicity testing would identify the exact concentrations at which a chemical exposure goes from having no effect (NOAEL) to causing structural malformations in susceptible

embryos. However, this is difficult to do *in vivo* due to the effects of maternal metabolism, chemical-specific pharmacokinetics, and a relatively limited window of teratogenic susceptibility. Historically, pregnant animals are treated relatively infrequently with high-dosages of the test chemical, and only a few different dosages are tested, which is problematic for several reasons. First, human exposures usually occur at chronic low-level doses, and it is likely that some teratogenic effects will be the result of the cumulative damage to target tissues or organs over time, which is not reflected in the dosing profile of *in vivo* protocols even with the same cumulative daily exposure (Scialli *et al.*, 2018). Additionally, different exposure levels of a teratogenic chemical are likely to activate different adverse outcome pathways (Ankley *et al.*, 2010). This effect was seen in the case of fluoxetine, which affects serotonin reuptake at therapeutic concentrations, inhibits canonical Wnt signaling at slightly higher concentrations and adversely affects cellular proliferation and viability via an unknown mechanism(s) at even higher concentrations. Third, high dosages like those used in animal models may cause fetal death and obscure the presence of malformations occurring at lower doses. “[D]ose metric provides the most meaningful basis for interpreting observed effects” of DART testing in animal models to a prediction of the effects in human pregnancies (Ankley *et al.*, 2010). Without establishing dose-response curves for teratogenic effects, it is risky to attempt to extrapolate results to an estimate of human risk. *In vitro* tests are well-suited to aid in the creation of dose-response curves because it is easy to alter doses to more accurately match human pharmacokinetic data and gene expression assays are often able to give an indication of dose-response before effects are seen at the organ or organism level. Therefore, discussion of dose is a perfect example of how *in vitro* testing methods can supplement, inform and enhance the power of *in vivo* testing.

Despite many advantages, there are limitations to *in vitro* screening that must be addressed. The robustness, flexibility and efficiency of *in vitro* tests largely come from their inherent simplicity. Rather than trying to understand drug effects on multiple organ systems, *in vitro* screens assess effects on a single cell type. As a result, it is unlikely that any single *in vitro* assay will be able to fully represent the broad spectrum of teratogenic susceptibility during development. To address this, coordinated panels of *in vitro* models, each representing different developmental stages or cell types, will likely be used together to screen for teratogenic chemicals. Even so, there is a limited ability of *in vitro* cells to represent the collective responses of tissues within a complex organism, which maintains homeostasis through several compensatory mechanisms. The best example of this is the effect of maternal metabolism on the actual embryonic exposure to developmental toxicants. Although it may be possible to simulate maternal metabolism *in vitro*, that and other aspects of biologic adaptability remain a hurdle for *in vitro* screens (Luijten *et al.*, 2008). Additionally, a model is only valuable to the extent that it can simulate the process or system that we wish to predict. To accurately interpret the results of *in vitro* screens, we must understand what embryologic processes are represented and ensure that those processes are a valid reflection of human biology or development. Many *in vitro* models attempt to mimic a complex developmental process

in the absence of the three-dimensional structural complexity, patterning, spatial dynamics and cellular interactions that are ubiquitous in organismal development (Tonk *et al.*, 2015). This issue was also discussed by Scialli *et al.* as well, as they wrote, “*When modeling developmental processes and the toxicity that disrupts them, we need to rebuild this complexity*” (Scialli *et al.*, 2018). The balance between quantifiable simplicity and realistic complexity will likely be addressed through the coordination of a variety of *in vitro* models and advancements in the way developmental toxicology data is accumulated and processed (De Abrew *et al.*, 2016; Ankley *et al.*, 2010; Knudsen *et al.*, 2011). If we remember the limitations of *in vitro* models and interpret results accordingly, these alternative developmental toxicity assays are a powerful tool in the study of teratology and embryonic development.

## **5.2 - The P19C5 morphogenesis model adds complexity to *in vitro* developmental toxicity screens**

P19C5 model recreates some of the structural complexity and cellular heterogeneity that many *in vitro* models lack. After the initial induction of differentiation, P19C5 EBs do not require any additional experimental manipulations to undergo *in vitro* processes similar to embryonic gastrulation and morphogenesis. Since EBs demonstrate a substantial ability to self-organize and establish body axes and patterning in a consistent temporal manner, it is likely that they are following the early progression of developmental steps that create embryonic complexity *in vivo*. The process of morphogenesis represents the successful function of many critical processes, so the consistent Day 4 elongation of EBs further supports that prerequisite developmental processes are also represented in the P19C5 embryoid body model. A biologically active chemical can exert effects in three general manners: 1) acting against a single specific target or pathway, 2) non-selectively disrupting many pathways or targets (e.g., chemicals that have general toxicity), and 3) affecting multiple, unrelated biologic pathways partially selective manner. In making recommendations for the field of predictive toxicology, Daston *et al.* (2015) assert that this third type of chemical is the most difficult to using *in vitro* systems and that, “[s]uch chemicals with pleiotropic behavior will probably require more sophisticated systems biology approaches to predict which types of toxicity might be anticipated, and at what concentrations” (Daston *et al.*, 2015). Because P19C5 EBs demonstrate a higher level of complexity than most *in vitro* models, they may be better suited to identify the teratogenic chemicals that have pleiotropic effects and can disrupt several developmental pathways. Finally, the P19C5 model is useful because it can be used for high-content genomic assays and may be amenable to modifications allowing high-throughput screening. Because the P19C5 EBs are simple to create and are evaluated using simple morphologic parameters that can be assessed with largely automated image analysis software, it is theoretically possible to automate many parts of the P19C5 model in a way that would allow high-throughput screening. If P19C5 EBs are to be used in regulatory testing, we also must consider the limitations of the model. P19C5 cells are a murine stem cell line, and like *in vivo* rodent models, the P19C5 cells do not respond to primate-specific teratogens like

thalidomide. In the future, we would like to create a similar morphogenesis-based model using human cell lines, which would respond to these species-specific developmental toxicants. Additionally, P19C5 EBs develop and elongate without detectable Shh signaling activity, which is crucial for normal development of anterior structures. This may be because P19C5 EBs appear to recreate only the caudal portion of embryonic gastrulation and elongation morphogenesis. If so, it is appealing to hypothesize that the high activity of caudal signaling pathways such as RA, Fgf and Wnt may suppress Shh activity. Nonetheless, it is important to consider the advantages as well as the limitations of the P19C5 model when it is applied to developmental toxicity research.

### 5.3 - Future directions and concluding remarks

Within the past decade, the focus in developmental toxicology has shifted away from animal-based testing methods towards *in vitro* methods. Animal-based toxicity screens are an unwieldy, low-throughput approach that cannot reasonably meet the need to comprehensively screen high numbers of chemicals for potential developmental toxicity. In 2007, the National Research Council released the report, "Toxicity Testing in the 21<sup>st</sup> Century," (TT21C) which emphasized the need for a paradigm shift in regulatory toxicology towards an approach based on *in vitro* methods capable of high-throughput screening of chemicals rather continuing the traditional reliance on resource-intensive *in vivo* models (Committee on Toxicity Testing and Assessment of Environmental Agents, 2007). The TT21C report described the value of assessing teratogenic mechanisms and discussed the concept of "toxicity pathways" as a framework for the study of teratogens. The TT21C report was followed by the Organization for Economic Co-operation and Development (OECD) report in 2012 that promoted adverse outcome pathways (AOPs) as a more informative and relevant structure for future DART research and screening. Although the terminology varies slightly, the NRC toxicity pathways and the OECD adverse outcome pathways, represent similar functional paradigms that emphasize the importance of identifying the molecular initiation event (MIE) for any given teratogen. Ankley *et al.* (2010) defined the AOP as, "*a conceptual construct that portrays existing knowledge concerning the linkage between a direct molecular initiating event and an adverse outcome at a biological level of organization relevant to risk assessment...[and] can focus toxicity testing in terms of species and endpoint selection, enhance across-chemical extrapolation, and ... facilitate use of molecular or biochemical endpoints (sometimes referred to as biomarkers) for forecasting chemical impacts on individuals and populations.*" (Ankley *et al.*, 2010). The AOP provides a structure for mechanistic information from many types of toxicologic models to be integrated into a single cohesive description of teratogenic mechanism.

Compared to *in vivo* DART testing, the AOP method of DART assessment allows more inclusive incorporation of existing knowledge on teratogenic mechanisms, a better identification of endpoints relevant to assessing human risk, and more efficient delegation of limited resources for DART testing to the areas where additional information on teratogenic mode of action is vital to risk assessment (Ankley *et*

*al.*, 2010). If testing is done at the mechanistic (or “mode of action” level), then it may be possible to identify “critical biological targets” (Daston *et al.*, 2015) that will enable the identification of “a small number of key events at a molecular or cellular level that predict an adverse outcome for which testing could be preformed *in vitro* or *in silico* or, rarely, using limited *in vivo* models” (Scialli *et al.*, 2018). The future of developmental toxicology research will emphasize the importance of the third of Wilson’s principles: “Developmental toxins act in specific ways (mechanisms) on developing cells and tissues to initiate sequences of abnormal developmental events (pathogenesis)” (Faqi *et al.*, 2012). This represents a hypothesis-based approach to teratology research that seeks to answer key questions rather than blindly screening chemicals using outdated methods. *“A key component in the development of hypothesis-driven testing is the understanding that there are a finite number of modes of action involved in developmental toxicity. We do not yet know all possible modes of action or the number of pathways that could be involved, but we believe that these pathways are knowable”* (Scialli *et al.*, 2018). The AOP framework allows the integration of information from several *in vitro* models to create a comprehensive picture of potential developmental toxicity for a given chemical that can be more easily extrapolated to a risk assessment in human pregnancies. The P19C5 model and other *in vitro* systems are a promising addition to the field of developmental toxicology research and may one day be a fundamental part of regulatory DART testing for pharmaceutical chemicals.

## APPENDIX A: MACRO SCRIPT FOR IMAGEJ MORPHOMETRIC ANALYSIS OF EBs

I wrote the following macro code to semi-automate the morphometric analysis of P19C5 EBs. This ImageJ macro (written in the ImageJ language, “.ijm,”) enabled EB images to be analyzed in approximately one tenth of the the amount of time that it took to outline EBs by hand, using the polygon selection tool in ImageJ. In order for the P19C5 system to accommodate high-throughput analyses, it must be amenable to automation and modifications that increase efficiency. Since a major drawback of both *in vivo* testing and the ESTc is the necessity for skilled, semi-quantitative scoring of apical endpoints, this macro demonstrates that the P19C5 system may be an advantageous innovation for high-throughput screening in DART testing.

Yellow highlighted regions within the following code indicate command sequences that are illustrated with pictures in Appendix B. The bold, red letters corresponding to the highlighted phrases indicate which picture matches the text and are not part of the macro code. The macro code is as follows:

```
roiManager("reset");
(A) path = File.openDialog("Select file");
open(path);

(B) Dialog.create("EB Image Analysis Parameters");
Dialog.addNumber("Analyze_Particles_(Min):", 10000);
Dialog.addNumber("Analyze_Particles_(Max):", 10000000);
Dialog.addNumber("Interpolation Interval:", 1);
Dialog.show();

APmin = Dialog.getNumber();
APmax = Dialog.getNumber();
Interval = Dialog.getNumber();

name=File.nameWithoutExtension;
parent=File.getParent(path);
(C) run("Set Scale...", "distance=0.3940 known=1 pixel=1 unit=uM global");

run("Duplicate...", name+"-1");
(D) run("Invert");
(E) imageCalculator("Min create", name+"-1.tif", name+".tif");
selectWindow("Result of "+name+"-1.tif");

(F) run("Find Edges");
run("8-bit");
(G) setMinAndMax(0, 125);
```

```

call("ij.ImagePlus.setDefault16bitRange", 8);
run("Apply LUT");
setMinAndMax(40, 125);
call("ij.ImagePlus.setDefault16bitRange", 8);
run("Apply LUT");

(H) //setTool("brush");
    setLineWidth(3);
    setForegroundColor(0,0,0);
    setBackgroundColor(255,255,255);

(I) waitForUser("Pause","Separate EBs with Paintbrush tool (Color = Black,
    Brush Width = 3 pixels). Click OK when done.");

(J) run("Make Binary");

    for(i=0;i<3;i++){
(K)     run("Close-");
(L)     run("Fill Holes");
    }

    waitForUser("Pause","Are all EBs complete and/or separated? Click OK if
    yes.");

(M) run("Erode");

run("Analyze Particles...", "size=&APmin-&APmax clear add");

roiManager("Show All with labels");
for(i=0; i<roiManager("count"); i++){
    roiManager("select", i);
    run("Interpolate", "interval=&Interval");
(N)     run("Fit Spline");
        roiManager("Update");
    }

(O) run("Set Measurements...", "area mean standard modal min centroid center
    perimeter bounding fit shape feret's integrated median skewness kurtosis
    area_fraction display add redirect=None decimal=3");
roiManager("Deselect");
selectWindow(name+".tif");
roiManager("Show All with labels");
(P) roiManager("Measure");

```

```
selectWindow(name+".tif");
save(path);

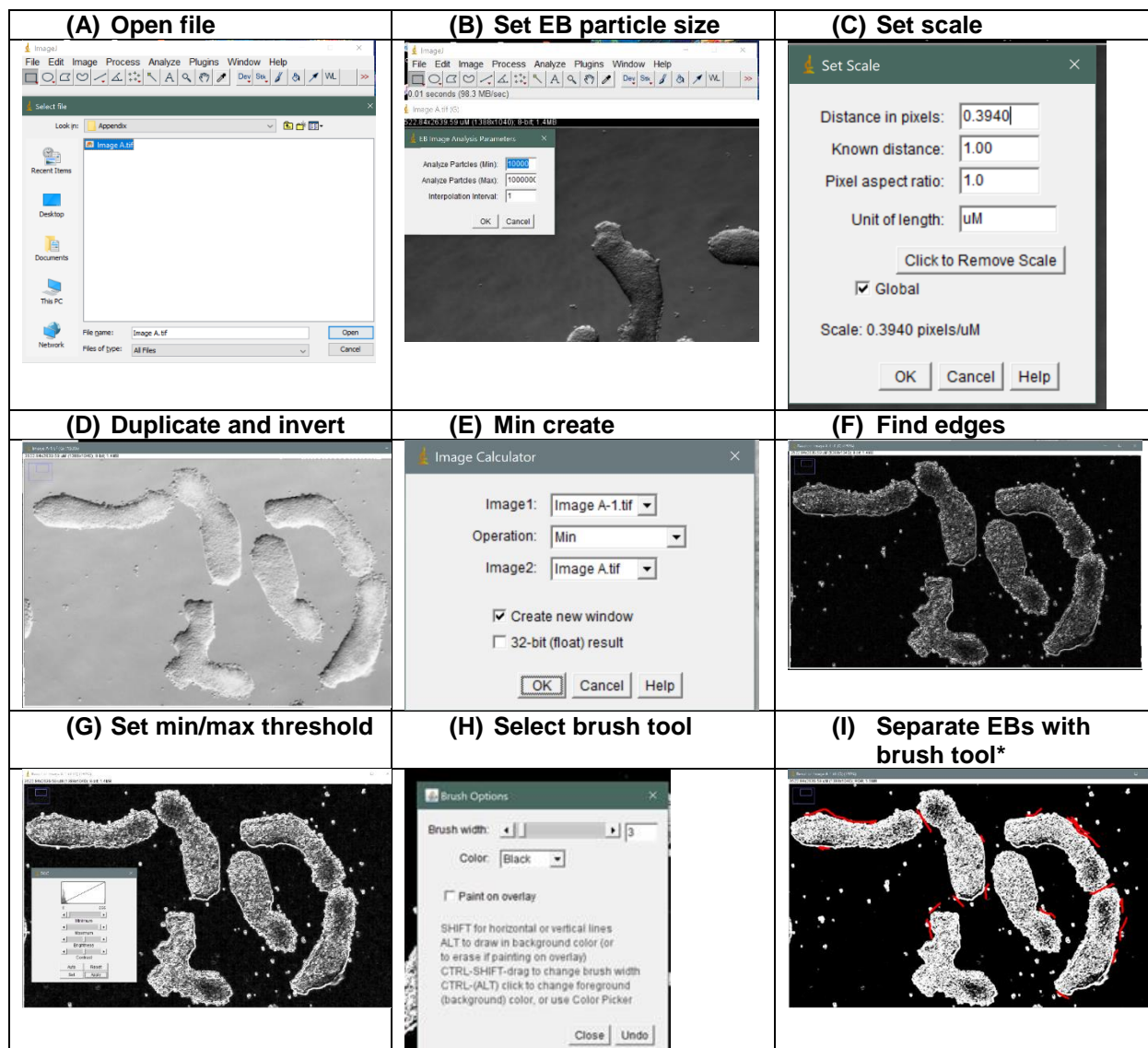
output=getDirectory("Choose Folder for Results");
selectWindow("Results");
saveAs("Results", output+name+"_results.csv");

selectWindow(name+"-1.tif");
close();
selectWindow("Result of "+name+"-1.tif");
close();
```

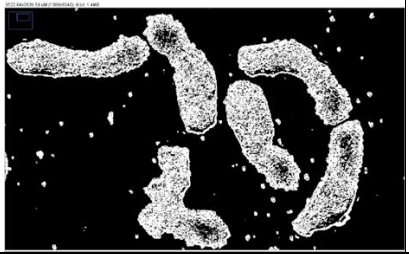


## APPENDIX B: EXAMPLE IMAGES CORRESPONDING TO APPENDIX A MACRO SCRIPT FOR IMAGEJ MORPHOMETRIC ANALYSIS OF EBS


Images to accompany EB image analysis macro script are demonstrative of the general concepts and .ijm commands only. Different ImageJ versions may look slightly different. Some variables of the macro script (i.e., particle size, min/max threshold, etc.) may need to be adjusted if using this macro for images captured through any methods other than those described in the materials and methods portions of the preceding text.




(J) Make binary




(K) Close



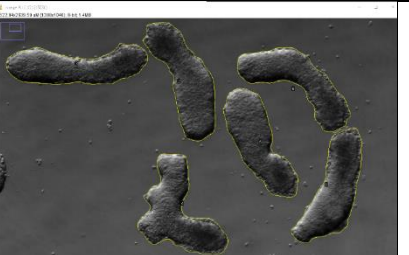
(L) Fill holes



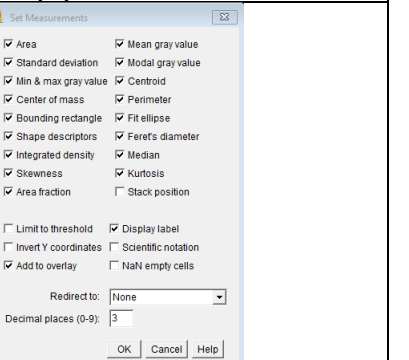
(M) Erode



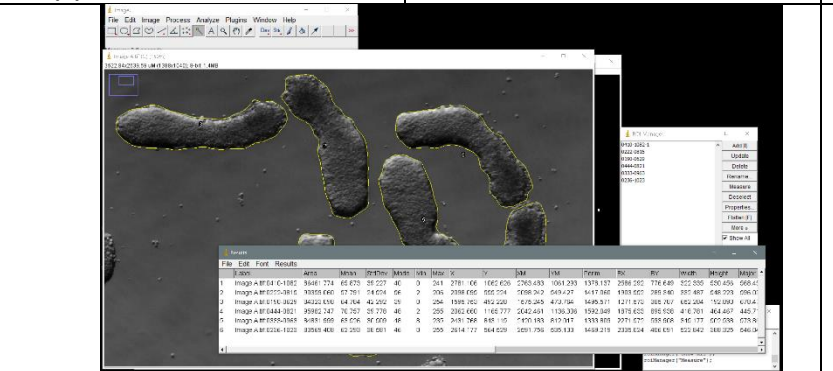
(N) Interpolate/fit spline



(O) Set measurements



(P) Measure



\* Note: Brush lines in (I) were illustrated in red. During actual image analysis brush lines would be black.

## REFERENCES

- De Abrew, K. N., Kainkaryam, R. M., Shan, Y. K., Overmann, G. J., Settivari, R. S., Wang, X., Xu, J., Adams, R. L., Tiesman, J. P., Carney, E. W., Naciff, J. M., and Daston, G. P. (2016). Grouping 34 chemicals based on mode of action using connectivity mapping. *Toxicol. Sci.*, **151**, 447–461.
- Abu-Heija, A. T., Fleming, R., Yates, R. W., and Coutts, J. R. (1995). Pregnancy outcome following exposure to gonadotrophin-releasing hormone analogue during early pregnancy: comparisons in patients with normal or elevated luteinizing hormone. *Hum. Reprod.* **10**, 3317–3319.
- Ai, D., Fu, X., Wang, J., Lu, M. F., Chen, L., Baldini, A., Klein, W. H., and Martin, J. F. (2007). Canonical Wnt signaling functions in second heart field to promote right ventricular growth. *Proc. Natl. Acad. Sci.*, **104**, 9319–9324.
- Al-Qattan, M. M., Al-Balwi, M., Eyaid, W., Al-Abdulkarim, I., and Al-Turki, S. (2009). Congenital duplication of the palm syndrome: gene analysis and the molecular basis of its clinical features. *J. Hand. Surg. Eur. Vol.* **34**, 247–251.
- Amsterdam, J. D., Fawcett, Jan Quitkin, Frederic M., Reimherr, Frederick W., Rosenbaum, Jerrold F., Michelson, David Hornig-Rohan, Mady Beasley, Charles M. (1997). Fluoxetine and norfluoxetine plasma concentrations in major depression: A multicenter study. *Am. J. Psychiatry*, **154**, 963–969.
- Andrews, J. E., Ebron-McCoy, M., Kavlock, R. J., and Rogers, J. M. (1995). Developmental toxicity of formate and formic acid in whole embryo culture: a comparative study with mouse and rat embryos. *Teratology* **51**, 243–251.
- Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., Mount, D. R., Nichols, J. W., Russom, C. L., Schmieder, P. K., Serrano, J. A., Tietge, J. E., and Villeneuve, D. L. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.*, **29**, 730–741.
- Aubert, N., Ameller, T., and Legrand, J. J. (2012). Systemic exposure to parabens: Pharmacokinetics, tissue distribution, excretion balance and plasma metabolites of [14C]-methyl-, propyl- and butylparaben in rats after oral, topical or subcutaneous administration. *Food Chem. Toxicol.*, **50**, 445–454.
- Azhar, M., and Ware, S. M. (2016). Genetic and Developmental Basis of Cardiovascular Malformations. *Clin. Perinatol.*, **43**, 39–53.
- Bailey, J., Knight, A., and Balcombe, J. (2005). The future of teratology research is in vitro. *Biogenic. Amines.* **19**, 97–145.

Bain, G., Ray, W. J., Yao, M., and Gottlieb, D. I. (1994). From embryonal carcinoma cells to neurons: the P19 pathway. *Bioessays* **16**, 343–348.

Barbaric, I., Miller, G., and Dear, T. N. (2007). Appearances can be deceiving: phenotypes of knockout mice. *Brief. Funct. Genomic. Proteomic.* **6**, 91–103.

Barbero, P., Lotersztejn, V., Bronberg, R., Perez, M., and Alba, L. (2004). Acitretin embryopathy: a case report. *Birth Defects Res. A Clin. Mol. Teratol.* **70**, 831–833.

Bérard, A., and Kori, S. (2012). Dihydroergotamine (DHE). use during gestation and the risk of adverse pregnancy outcomes. *Headache.* **52**, 1085–1093.

Blanch, G., Quenby, S., Ballantyne, E. S., Gosden, C. M., Neilson, J. P., and Holland, K. (1998). Embryonic abnormalities at medical termination of pregnancy with mifepristone and misoprostol during first trimester: observational study. *BMJ.* **316**, 1712–1713.

Blankart, C. R., Stargardt, T., and Schreyögg, J. (2011). Availability of and access to orphan drugs: An international comparison of pharmaceutical treatments for pulmonary arterial hypertension, Fabry disease, hereditary angioedema and chronic myeloid leukaemia. *Pharmacoeconomics*, **29**, 63–82.

Bolo, N. R., Hodé, Y., Nédélec, J. F., Lainé, E., Wagner, G., and Macher, J. P. (2000). Brain pharmacokinetics and tissue distribution in vivo of fluvoxamine and fluoxetine by fluorine magnetic resonance spectroscopy. *Neuropsychopharmacology.* **23**, 428–438.

Bos-Thompson, M. A., Hillaire-Buys, D., Roux, C., Faillie, J. L., and Amram, D. (2008). Möbius syndrome in a neonate after mifepristone and misoprostol elective abortion failure. *Ann. Pharmacother.* **42**, 888–892.

Bowie, M., Pilie, P., Wulfkühle, J., Lem, S., Hoffman, A., Desai, S., Petricoin, E., Carter, A., Ambrose, A., Seewaldt, V., Yu, D., and Ibarra-Drendall, C. (2015). Fluoxetine induces cytotoxic endoplasmic reticulum stress and autophagy in triple negative breast cancer. *World J. Clin. Oncol.* **6**, 299–311.

Brent, R. L. (1964). Drug testing in animals for teratogenic effects. Thalidomide in the pregnant rat. *J. Pediatr.* **64**, 762–770.

van den Brink, S. C., Baillie-Johnson, P., Balayo, T., Hadjantonakis, A. K., Nowotschin, S., Turner, D. A., and Martinez-Arias, A. (2014). Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Development* **141**, 4231–4242.

- Brown, N. A. (2002). Selection of test chemicals for the ECVAM international validation study on in vitro embryotoxicity tests. European Centre for the Validation of Alternative Methods. *Altern. Lab Anim.* **30**, 177–198.
- Browne, C. D., Hindmarsh, E. J., and Smith, J. W. (2006). Inhibition of endothelial cell proliferation and angiogenesis by orlistat, a fatty acid synthase inhibitor. *FASEB J.* **20**, 2027–2035.
- Brunswick, D. J., Amsterdam, J. D., Fawcett, J., Quitkin, F. M., Reimherr, F. W., Rosenbaum, J. F., and Beasley, C. M. Jr. (2002). Fluoxetine and norfluoxetine plasma concentrations during relapse-prevention treatment. *J. Affect. Disord.* **68**, 243–249.
- Buesen, R., Visan, A., Genschow, E., Slawik, B., Spielmann, H., and Seiler, A. (2004). Trends in improving the embryonic stem cell test (EST): an overview. *ALTEX* **21**, 15–22.
- Buesen, R., Genschow, E., Slawik, B., Visan, A., Spielmann, H., Luch, A., and 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.
- Chaabane S, Bérard A. (2013). Epidemiology of major congenital malformations with specific focus on teratogens. *Curr. Drug Saf.* **8**, 128–140.
- Cahill, D. J., Fountain, S. A., Fox, R., Fleming, C. F., Brinsden, P. R., and Hull, M. G. (1994). Outcome of inadvertent administration of a gonadotrophin-releasing hormone agonist (buserelin). in early pregnancy. *Hum. Reprod.* **9**, 1243–1246.
- Charles, D. H., Ness, A. R., Campbell, D., Smith, G. D., Whitley, E., and Hall, M. H. (2005). Folic acid supplements in pregnancy and birth outcome: re-analysis of a large randomised controlled trial and update of Cochrane review. *Paediatr. Perinat. Epidemiol.* **19**, 112–124.
- Chen, J., Wang, J., Shao, J., Gao, Y., Xu, J., Yu, S., Liu, Z., and Jia, L. (2014). The unique pharmacological characteristics of mifepristone (RU486): from terminating pregnancy to preventing cancer metastasis. *Med. Res. Rev.* **34**, 979–1000.
- Cheer, S. M., and Goa, K. L. (2001). Fluoxetine: a review of its therapeutic potential in the treatment of depression associated with physical illness. *Drugs.* **61**, 81–110.
- Choi, J. H., Jeong, Y. J., Yu, A. R., Yoon, K. S., Choe, W., Ha, J., Kim, S. S., Yeo, E. J., and Kang, I. (2017). Fluoxetine induces apoptosis through endoplasmic reticulum stress via mitogen-activated protein kinase activation and histone hyperacetylation in SK-N-BE(2)-M17 human neuroblastoma cells. *Apoptosis.* **22**, 1079–1097.

- Clagett-Dame, M., and Knutson, D. (2011). Vitamin A in reproduction and development. *Nutrients*, **3**, 385–428.
- Clark, R. L., Lerman, S. A., Cox, E. M., Gristwood, W. E., and White, T. E. K. (2008). Developmental toxicity of artesunate in the rat: Comparison to other artemisinins, comparison of embryotoxicity and kinetics by oral and intravenous routes, and relationship to maternal reticulocyte count. *Birth Defects Res. B Dev. Reprod. Toxicol.* **83**, 397–406.
- Clark, R. L. (2009). Embryotoxicity of the artemisinin antimalarials and potential consequences for use in women in the first trimester. *Reprod. Toxicol.*, **28**, 285–296.
- Cockshott, I. D. (2004). Bicalutamide: clinical pharmacokinetics and metabolism. *Clin. Pharmacokinet.* **43**, 855–878.
- Committee on Toxicity Testing and Assessment of Environmental Agents, N. R. C. (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy The National Academies Press, Washington D. C.
- Copp, A. J., Greene, N. D., and Murdoch, J. N. (2003). Dishevelled: linking convergent extension with neural tube closure. *Trends Neurosci.* **26**, 453–455.
- van Dartel, D. A., and Piersma, A. H. (2011). The embryonic stem cell test combined with toxicogenomics as an alternative testing model for the assessment of developmental toxicity. *Reprod. Toxicol.* **32**, 235–244.
- Daston, G. P. (2011). Laboratory models and their role in assessing teratogenesis. *Am. J. Med. Genet. C Semin. Med. Genet.* **157C**, 183–187.
- Daston, G., Knight, D. J., Schwarz, M., Gocht, T., Thomas, R. S., Mahony, C., and Whelan, M. (2015). SEURAT: Safety Evaluation Ultimately Replacing Animal Testing--recommendations for future research in the field of predictive toxicology. *Arch. Toxicol.*, **89**, 15–23.
- Daston, G. P., Chapin, R. E., Scialli, A. R., Piersma, A. H., Carney, E. W., Rogers, J. M., and Friedman, J. M. (2010). A different approach to validating screening assays for developmental toxicity. *Birth Defects Res. B Dev. Reprod. Toxicol.* **89**, 526–530.
- Daston, G. P., Beyer, B. K., Carney, E. W., Chapin, R. E., Friedman, J. M., Piersma, A. H., Rogers, J. M., and Scialli, A. R. (2014). Exposure-based validation list for developmental toxicity screening assays. *Birth Defects Res. B Dev. Reprod. Toxicol.* **101**, 423–428.

- De Santis, M., Carducci, B., Cavaliere, A. F., De Santis, L., Lucchese, A., Straface, G., and Caruso, A. (2003). Paternal exposure to ribavirin: pregnancy and neonatal outcome. *Antivir. Ther.* **8**, 73–75.
- Diav-Citrin, O., Shechtman, S., Weinbaum, D., Wajnberg, R., Avgil, M., Di Gianantonio, E., Clementi, M., Weber-Schoendorfer, C., Schaefer, C., and Ornoy, A. (2008). Paroxetine and fluoxetine in pregnancy: a prospective, multicentre, controlled, observational study. *Br. J. Clin. Pharmacol.* **66**, 695-705.
- Ellfolk, M., and Malm, H. (2010). Risks associated with in utero and lactation exposure to selective serotonin reuptake inhibitors (SSRIs). *Reprod. Toxicol.* **30**, 249-260.
- Elmazar, M. M., and Nau, H. (2004). Potentiation of the teratogenic effects induced by coadministration of retinoic acid or phytanic acid/phytol with synthetic retinoid receptor ligands. *Arch. Toxicol.* **78**, 660–668.
- Eskes, S. A., and Wiersinga, W. M. (2009). Amiodarone and thyroid. *Best Pract. Res. Clin. Endocrinol. Metab.* **23**, 735–751.
- Estevan, C., Romero, A. C., Pamies, D., Vilanova, E., and Sogorb, M. A. (2009). Embryonic Stem Cells in Toxicological Studies.
- Faqi, A. S., Hoberman, A., Lewis, E., and Stump, D. (2012). Developmental and Reproductive Toxicology. In, Faqi, A. S. (ed), *A Comprehensive Guide to Toxicology in Preclinical Drug Development*. Elsevier Science & Technology, pp. 335–364.
- Farr, S. L., Bitsko, R. H., Hayes, D. K., and Dietz, P. M. (2010). Mental health and access to services among US women of reproductive age. *Am. J. Obstet. Gynecol.* **203**, 542. e1-9.
- FDA. (2001). *Clarinx pharmacology review*. Retrieved from [http://www.accessdata.fda.gov/drugsatfda\\_docs/nda/2013/021187Orig1s021.pdf](http://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/021187Orig1s021.pdf)
- FDA. (2010). *Gilenya Pharmacology Review (Application number: 22-527)*. Retrieved from [http://www.accessdata.fda.gov/drugsatfda\\_docs/nda/2010/022527Orig1s000TO.cfm](http://www.accessdata.fda.gov/drugsatfda_docs/nda/2010/022527Orig1s000TO.cfm)
- Franco, P. (2013). Orphan drugs: The regulatory environment. *Drug Discov. Today*, **18**, 163–172.
- Friedman, J. M. (2009). Big risks in small groups: The difference between epidemiology and counselling. *Birth Defects Res. A Clin. Mol. Teratol.* **85**, 720–724.
- Friedman, J. M. (2010). The principles of teratology: are they still true? *Birth Defects Res. A Clin. Mol. Teratol.* **88**, 766–768.

- Gammie, T. Gammie, T., Lu, C. Y., and Ud-Din Babar, Z. (2015). Access to orphan drugs: A comprehensive review of legislations, regulations and policies in 35 countries. *PLoS One*, **10**, 1–24.
- Gao, S. Y., Wu, Q. J., Zhang, T. N., Shen, Z. Q., Liu, C. X., Xu, X., Ji, C., and Zhao, Y. H. (2017). Fluoxetine and congenital malformations: a systematic review and meta-analysis of cohort studies. *Br. J. Clin. Pharmacol.* **83**, 2134-2147.
- Gao, X., Yourick, J. J., and Sprando, R. L. (2014). Transcriptomic characterization of C57BL/6 mouse embryonic stem cell differentiation and its modulation by developmental toxicants. *PLoS One* **9**, e108510.
- Geiger, J. M., Baudin, M., and Saurat, J. H. (1994). Teratogenic risk with etretinate and acitretin treatment. *Dermatology*. **189**, 109–116.
- Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N., Piersma, A., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., *et al.* (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.
- Gessert, S., and Kühl, M. (2010). The multiple phases and faces of wnt signaling during cardiac differentiation and development. *Circ. Res.* **107**, 186-199.
- Gilbert, S. F. (2010). *Developmental Biology* 9th ed. Sinauer Associates, Inc., Sunderland, MA.
- Gilboa, S. M., Ailes, E. C., Rai, R. P., Anderson, J. A., and Honein, M. A. (2014). Antihistamines and birth defects: a systematic review of the literature. *Expert. Opin. Drug Saf.* **13**, 1667–1698.
- Gizzo, S., Saccardi, C., Patrelli, T. S., Berretta, R., Capobianco, G., Di Gangi, S., Vacilotto, A., Bertocco, A., Noventa, M., Ancona, E., *et al.* (2013). Update on raloxifene: mechanism of action, clinical efficacy, adverse effects, and contraindications. *Obstet. Gynecol. Surv.* **68**, 467–481.
- Gocht, T., Berggren, E., Ahr, H. J., Cotgreave, I., Cronin, M. T. D., Daston, G., Hardy, B., Heinzle, E., Hescheler, J., Knight, D. J., Mahony, C., Peschanski, M., Schwarz, M., Thomas, R. S., Verfaillie, C., White, A., and Whelan, M. (2015). The SEURAT-1 approach towards animal free human safety assessment. *ALTEX*, **32**, 9–24.
- Godfrey, L. M., Erramouspe, J., and Cleveland, K. W. (2012). Teratogenic risk of statins in pregnancy. *Ann. Pharmacother.* **46**, 1419–1424.
- Gram, L. (1994). Fluoxetine. *N. Engl. J. Med.* **331**, 1354-1361.



Grigoryan, T., Wend, P., Klaus, A., and Birchmeier, W. (2008). Deciphering the function of canonical Wnt signals in development and disease: conditional loss- and gain-of-function mutations of beta-catenin in mice. *Genes Dev.* **22**, 2308-2241.

Hajihosseini, M. K. (2008). Fibroblast growth factor signaling in cranial suture development and pathogenesis. *Front. Oral. Biol.* **12**, 160–177.

Han, Y. S., and Lee, C. S. (2009). Antidepressants reveal differential effect against 1-methyl-4-phenylpyridinium toxicity in differentiated PC12 cells. *Eur. J. Pharmacol.* **604**, 36-44.

Hartung, T. (2011). From alternative methods to a new toxicology. *Eur. J. Pharm. Biopharm.* **77**, 338–349.

Hartung, T., and Rovida, C. (2009). Chemical regulators have overreached. *Nature.* **460**, 1080–1081.

Hayess, K., Riebeling, C., Pirow, R., Steinfath, M., Sittner, D., Slawik, B., Luch, A., and Seiler, A. E. (2013). The DNT-EST: a predictive embryonic stem cell-based assay for developmental neurotoxicity testing in vitro. *Toxicology.* **314**, 135–147.

Herion, N. J. Herion, N. J., Salbaum, J. M., and Kappen, C. (2014). Traffic jam in the primitive streak: The role of defective mesoderm migration in birth defects. *Birth Defects Res. Part A - Clin. Mol. Teratol.*, **100**, 608–622.

Hettwer, M., Reis-Fernandes, M. A., Iken, M., Ott, M., Steinberg, P., and Nau, H. (2010). Metabolic activation capacity by primary hepatocytes expands the applicability of the embryonic stem cell test as alternative to experimental animal testing. *Reprod. Toxicol.* **30**, 113–120.

van der Heyden, M. A., and Defize, L. H. (2003). Twenty one years of P19 cells: what an embryonal carcinoma cell line taught us about cardiomyocyte differentiation. *Cardiovasc. Res.* **58**, 292–302.

Hiemke, C., and Härtter, S. (2000). Pharmacokinetics of selective serotonin reuptake inhibitors. *Pharmacol. Ther.* **85**, 11-28.

Hohjoh, H. (2013). MicroRNA expression during neuronal differentiation of human teratocarcinoma NTera2D1 and mouse embryonic carcinoma P19 cells. *Methods Mol. Biol.* **936**, 257–269.

Hohmann, M., and Künzel, W. (1992). Dihydroergotamine causes fetal growth retardation in guinea pigs. *Arch. Gynecol. Obstet.* **251**, 187–192.

- Huang, S. M., Mishina, Y. M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G. A., Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., *et al.* (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614-620.
- Hui, J., Zhang, J., Kim, H., Tong, C., Ying, Q., Li, Z., Mao, X., Shi, G., Yan, J., Zhang, Z., and Xi, G. (2015). Fluoxetine regulates neurogenesis in vitro through modulation of GSK-3 $\beta$ / $\beta$ -catenin signaling. *Int. J. Neuropsychopharmacol.* **18**, pii: pyu099.
- Ito, T., and Handa, H. (2012). Deciphering the mystery of thalidomide teratogenicity. *Congenit. Anom.* **52**, 1–7.
- 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.
- Jacqz-Aigrain, E., and Koren, G. (2005). Effects of drugs on the fetus. *Semin. Fetal Neonatal Med.*, **10**, 139–147.
- Jelínek, R. (2005). The contribution of new findings and ideas to the old principles of teratology. *Reprod. Toxicol.* **20**, 295–300.
- Jones-Villeneuve, E. M., McBurney, M. W., Rogers, K. A., and Kalnins, V. I. (1982). Retinoic acid induces embryonal carcinoma cells to differentiate into neurons and glial cells. *J. Cell Biol.* **94**, 253–262.
- de Jong, E., van Beek, L., and Piersma, A. H. (2012). Osteoblast differentiation of murine embryonic stem cells as a model to study the embryotoxic effect of compounds. *Toxicol In Vitro.* **26**, 970–978.
- Källén, B. A. (2014). Antiobesity drugs in early pregnancy and congenital malformations in the offspring. *Obes. Res. Clin. Pract.* **8**, e571–e576.
- Kameoka, S., Babiarz, J., Kolaja, K., and Chiao, E. (2014). A high-throughput screen for teratogens using human pluripotent stem cells. *Toxicol. Sci.* **2137**, 76–90.
- Karson, C. N., Newton, J. E., Livingston, R., Jolly, J. B., Cooper, T. B., Sprigg, J., and Komoroski, R. A. (1993). Human brain fluoxetine concentrations. *J. Neuropsychiatry Clin. Neurosci.* **5**, 322-329.
- Kavlock, R., Chandler, K., Houck, K., Hunter, S., Judson, R., Kleinstreuer, N., Knudsen, T., Martin, M., Padilla, S., Reif, D., *et al.* (2012). Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chem. Res. Toxicol.*, **25**, 1287–1302.

- Kim, J., Riggs, K. W., Misri, S., Kent, N., Oberlander, T. F., Grunau, R. E., Fitzgerald, C., and Rurak, D. W. (2006). Stereoselective disposition of fluoxetine and norfluoxetine during pregnancy and breast-feeding. *Br. J. Clin. Pharmacol.* **61**, 155-163.
- Kim, J., Riggs, K. W., and Rurak, D. W. (2004). Stereoselective pharmacokinetics of fluoxetine and norfluoxetine enantiomers in pregnant sheep. *Drug Metab Dispos.* **32**, 212-221.
- King, C. T., Rogers, P. D., Cleary, J. D., and Chapman, S. W. (1998). Antifungal therapy during pregnancy. *Clin. Infect. Dis.* **27**, 1151-1160.
- Kioussi, C., Briata, P., Baek, S. H., Rose, D. W., Hamblet, N. S., Herman, T., Ohgi, K. A., Lin, C., Gleiberman, A., Wang, J., *et al.* (2002). Identification of a Wnt/Dvl/beta-Catenin --> Pitx2 pathway mediating cell-type-specific proliferation during development. *Cell* **111**, 673-685.
- Klaus, A., Saga, Y., Taketo, M. M., Tzahor, E., and Birchmeier, W. (2007). Distinct roles of Wnt/beta-catenin and Bmp signaling during early cardiogenesis. *Proc. Natl. Acad. Sci. U S A.* **104**, 18531-18536.
- Ko, J. Y., Farr, S. L., Dietz, P. M., and Robbins, C. L. (2012). Depression and treatment among U. S. pregnant and nonpregnant women of reproductive age, 2005-2009. *J. Womens Health* **21**, 830-836.
- Komm, B. S., and Mirkin, S. (2014). An overview of current and emerging SERMs. *J. Steroid Biochem. Mol. Biol.* **143**, 207-222.
- Korinek, V., Barker, N., Morin, P. J., van Wichen, D., de Weger, R., Kinzler, K. W., Vogelstein, B., and Clevers, H. (1997). Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**, 1784-1787.
- Knudsen, T. B., Kavlock, R. J., Daston, G. P., Stedman, D., Hixon, M., and Kim, J. H. (2011). Developmental toxicity testing for safety assessment: New approaches and technologies. *Birth Defects Res. Part B - Dev. Reprod. Toxicol.*, **92**, 413-420.
- Kruegel, J., and Miosge, N. (2010). Basement membrane components are key players in specialized extracellular matrices. *Cell. Mol. Life Sci.*, **67**, 2879-2895.
- Kuske, B., Pulyanina, P. Y., and zur Nieden, N. I. (2012). Embryonic stem cell test: stem cell use in predicting developmental cardiotoxicity and osteotoxicity. *Methods Mol. Biol.* **889**, 147-179.
- Kuwagata, M., Takashima, H., and Nagao, T. (1998). A comparison of the in vivo and in vitro response of rat embryos to 5-fluorouracil. *J. Vet. Med. Sci.* **60**, 93-99.

Kwon, C., Arnold, J., Hsiao, E. C., Taketo, M. M., Conklin, B. R., and Srivastava, D. (2007). Canonical Wnt signaling is a positive regulator of mammalian cardiac progenitors. *Proc. Natl. Acad. Sci. U S A.* **104**, 10894-10899.

Lalani, S. R., and Belmont, J. W. (2014). Genetic and developmental basis of congenital cardiovascular malformations. *Principles of Developmental Genetics* **57**, 607-633.

Lankas, G. R., Cukierski, M. A., and Wise, L. D. (2004). The role of maternal toxicity in lovastatin-induced developmental toxicity. *Birth Defects Res. B Dev. Reprod. Toxicol.* **71**, 111–123.

Larson, J. L., Wallace, T. L., Tyl, R. W., Marr, M. C., Myers, C. B., and Cossum, P. A. (2000). The reproductive and developmental toxicity of the antifungal drug Nyotran (liposomal nystatin). in rats and rabbits. *Toxicol. Sci.* **53**, 421–429.

Lau, C. G., and Marikawa, Y. (2014). Morphology-based mammalian stem cell tests reveal potential developmental toxicity of donepezil. *Mol. Reprod. Dev.* **81**, 994–1008.

Lee, C. S., Kim, Y. J., Jang, E. R., Kim, W., and Myung, S. C. (2010). Fluoxetine induces apoptosis in ovarian carcinoma cell line OVCAR-3 through reactive oxygen species-dependent activation of nuclear factor-kappaB. *Basic Clin. Pharmacol. Toxicol.* **106**, 446-453.

Lee, H. Lee, H., Inselman, A. L. Kanungo, J., and Hansen, D. K. (2012). Alternative models in developmental toxicology. *Syst. Biol. Reprod. Med.*, **58**, 10–22.

Li, A. S. W., and Marikawa, Y. (2016). Adverse effect of valproic acid on an in vitro gastrulation model entails activation of retinoic acid signaling. *Reprod. Toxicol.* **66**, 68–83.

Li, A. S. W., and Marikawa, Y. (2015). An in vitro gastrulation model recapitulates the morphogenetic impact of pharmacological inhibitors of developmental signaling pathways. *Mol. Reprod. Dev.* **82**, 1015–1036.

Li,H., Rietjens, I. M., Louisse, J., Blok, M., Wang, X., Snijders, L., and van Ravenzwaay, B. (2015). Use of the ES-D3 cell differentiation assay, combined with the BeWo transport model, to predict relative in vivo developmental toxicity of antifungal compounds. *Toxicol. In Vitro.* **29**, 320–328.

Li, Q., Si, Y., Xie, L., Zhang, J., and Weina, P. (2009). Severe embryoletality of artesunate related to pharmacokinetics following intravenous and intramuscular doses in pregnant rats. *Birth Defects Res. B Dev. Reprod. Toxicol.* **86**, 385–393.

- Li, S. Li, S., Edgar, D., Fässler, R., Wadsworth, W., and Yurchenco, P. D. (2003). The role of laminin in embryonic cell polarization and tissue organization. *Dev. Cell*, **4**, 613–624.
- Liebner, S., Cattelino, A., Gallini, R., Rudini, N., Iurlaro, M., Piccolo, S., and Dejana, E. (2004). Beta-catenin is required for endothelial-mesenchymal transformation during heart cushion development in the mouse. *J. Cell Biol.* **166**, 359-367.
- Lin, L., Cui, L., Zhou, W., Dufort, D., Zhang, X., Cai, C. L., Bu, L., Yang, L., Martin, J., Kemler, R., *et al.* (2007). Beta-catenin directly regulates *Islet1* expression in cardiovascular progenitors and is required for multiple aspects of cardiogenesis. *Proc. Natl. Acad. Sci. U S A.* **104**, 9313-9318.
- Loebel, D. A. F., Watson, C. M., De Young, R. A., and Tam, P. P. L. (2003). Lineage choice and differentiation in mouse embryos and embryonic stem cells. *Dev. Biol.*, **264**, 1–14.
- Lu, H. C., Eichele, G., and Thaller, C. (1997). Ligand-bound RXR can mediate retinoid signal transduction during embryogenesis. *Development*. **124**, 195–203.
- Luijten, M., Verhoef, A., Westerman, A., and Piersma, A. H. (2008). Application of a metabolizing system as an adjunct to the rat whole embryo culture. *Toxicol. In Vitro* **22**, 1332–1336.
- Ma, L., Selamet Tierney, E. S., Lee, T., Lanzano, P., and Chung, W. K. (2012). Mutations in *ZIC3* and *ACVR2B* are a common cause of heterotaxy and associated cardiovascular anomalies. *Cardiol. Young.* **22**, 194–201
- Mahawong, P., Sinclair, A., Li, Y., Schlomer, B., Rodriguez, E. Jr., Ferretti, M. M., Liu, B., Baskin, L. S., and Cunha, G. R. (2014). Prenatal diethylstilbestrol induces malformation of the external genitalia of male and female mice and persistent second-generation developmental abnormalities of the external genitalia in two mouse strains. *Differentiation*. **88**, 51–69.
- Malm, H., Artama, M., Gissler, M., and Ritvanen, A. (2011). Selective serotonin reuptake inhibitors and risk for major congenital anomalies. *Obstet. Gynecol.* **118**, 111-120.
- Marikawa, Y., Tamashiro, D. A., Fujita, T. C., and Alarcon, V. B. (2009). Aggregated P19 mouse embryonal carcinoma cells as a simple in vitro model to study the molecular regulations of mesoderm formation and axial elongation morphogenesis. *Genesis*. **47**, 93–106.
- Marx-Stoelting, P., Adriaens, E., Ahr, H. J., Bremer, S., Garthoff, B., Gelbke, H. P., Piersma, A., Pellizzer, C., Reuter, U., Rogiers, V., *et al.* (2009). A review of the implementation of the embryonic stem cell test (EST). The report and recommendations of an ECVAM/ReProTect Workshop. *Altern. Lab. Anim.* **37**, 313–328.

- Marzocchi, M., and Lombardi, F. (2011). Dronedarone for atrial fibrillation therapy. *Expert. Rev. Cardiovasc. Ther.* **9**, 675–683.
- Matok, I., and Perlman, A. (2014). Metoclopramide in pregnancy: no association with adverse fetal and neonatal outcomes. *Evid. Based Med.* **19**, 115.
- McBurney, M. W. (1993). P19 embryonal carcinoma cells. *Int. J. Dev. Biol.* **37**, 135–140.
- McBurney, M. W., and Rogers, B. J. (1982). Isolation of male embryonal carcinoma cells and their chromosome replication patterns. *Dev. Biol.* **89**, 503–508.
- Mehndiratta, S., Suneja, A., Gupta, B., and Bhatt, S. (2010). Fetotoxicity of warfarin anticoagulation. *Arch. Gynecol. Obstet.* **282**, 335–337.
- van Mil, N. H., Oosterbaan, A. M., and Steegers-Theunissen, R. P. (2010). Teratogenicity and underlying mechanisms of homocysteine in animal models: a review. *Reprod. Toxicol.* **30**, 520–531.
- Miyamoto, K., Ohkawara, B., Ito, M., Masuda, A., Hirakawa, A., Sakai, T., Hiraiwa, H., Hamada, T., Ishiguro, N., and Ohno, K. (2017). Fluoxetine ameliorates cartilage degradation in osteoarthritis by inhibiting Wnt/ $\beta$ -catenin signaling. *PLoS One* **12**, e0184388.
- Morshed, K. M., Nagpaul, J. P., Majumdar, S., and Amma, M. K. P. (1988). Kinetics of propylene glycol elimination and metabolism in rat. *Biochem. Med. Metab. Biol.* **39**, 90–97.
- Murthy, R. K., Theriault, R. L., Barnett, C. M., Hodge, S., Ramirez, M. M., Milbourne, A., Rimes, S. A., Hortobagyi, G. N., Valero, V., and Litton, J. K. (2014). Outcomes of children exposed in utero to chemotherapy for breast cancer. *Breast Cancer Res.* **16**, 3414.
- Myles, N., Newall, H., Ward, H., and Large, M. (2013). Systematic meta-analysis of individual selective serotonin reuptake inhibitor medications and congenital malformations. *Aust. N. Z. J. Psychiatry.* **47**, 1002-1012.
- Narboux-Nême, N., Pavone, L. M., Avallone, L., Zhuang, X., and Gaspar, P. (2008). Serotonin transporter transgenic (SERT<sup>cre</sup>). mouse line reveals developmental targets of serotonin specific reuptake inhibitors (SSRIs). *Neuropharmacology* **55**, 994-1005.
- zur Nieden, N. I., Ruf, L. J., Kempka, G., Hildebrand, H., and Ahr, H. J. (2001). Molecular markers in embryonic stem cells. *Toxicol. In Vitro* **15**, 455–461.

Niederreither, K., Subbarayan, V., Dolle, P., and Chambon, P. (1999). Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet.* **21**, 444–448.

Niederreither, K., McCaffery, P., Dräger, U. C., Chambon, P., and Dollé, P. (1997). Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2). gene during mouse development. *Mech. Dev.*, **62**, 67–78.

Niederreither, K, V Subbarayan, P Dolle, and P Chambon. 1999. “Embryonic Retinoic Acid Synthesis Is Essential for Early Mouse Post-Implantation Development. ” *Nature Genetics* 21 (4): 444–48. doi:10.1038/7788., Niemann, S., Zhao, C., Pascu, F., Stahl, U., Aulepp, U., Niswander, L., Weber, J. L., and Müller, U. (2004). Homozygous WNT3 mutation causes tetra-amelia in a large consanguineous family. *Am. J. Hum. Genet.* 74, 558–563.

Niwa, H. (2007). How is pluripotency determined and maintained? *Development* **134**, 635–646.

Nowak, M. A., Boerlijst, M. C., Cooke, J., and Smith, J. M. (1997). Evolution of genetic redundancy. *Nature* **388**, 167–171.

Oglesby, L. A., Ebron, M. T., Beyer, P. E., Carver, B. D., and Kavlock, R. J. (1986). Co-culture of rat embryos and hepatocytes: in vitro detection of a proteratogen. *Teratog. Carcinog. Mutagen.* **6**, 129–138.

Ornoy, A., and Koren, G. (2017). Selective serotonin reuptake inhibitors during pregnancy: do we have now more definite answers related to prenatal exposure? *Birth Defects Res.* 109, 898-908.

Palmer, J. A., Smith, A. M., Egnash, L. A., Conard, K. R., West, P. R., Burrier, R. E., Donley, E. L. R., and Kirchner, F. R. (2013). Establishment and assessment of a new human embryonic stem cell-based biomarker assay for developmental toxicity screening. *Birth Defects Res. B Dev. Reprod. Toxicol.* **98**, 343–363.

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

Parker, S. E., Mai, C. T., Canfield, M. A., Rickard, R., Wang, Y., Meyer, R. E., Anderson, P., Mason, C. A., Collins, J. S., Kirby, R. S., Correa, A. ; National Birth Defects Prevention Network. (2010). Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004-2006. *Birth Defects Res. A Clin. Mol. Teratol.* **88**, 1008–1016.

Pasternak, B., Svanström, H., Mølgaard-Nielsen, D., Melbye, M., and Hviid, A. (2013). Metoclopramide in pregnancy and risk of major congenital malformations and fetal death. *JAMA.* **310**, 1601–1611.

Pennings, J. L. A., van Dartel, D. A. M., Robinson, J. F., Pronk, T. E., and Piersma, A. H. (2011). Gene set assembly for quantitative prediction of developmental toxicity in the embryonic stem cell test.

*Toxicology*, **284**, 63–71.

Piersma, A. H., Genschow, E., Verhoef, A., Spanjersberg, M. Q. I., Brown, N. A., Brady, M., Burns, A., Clemann, N., Seiler, A., and Spielmann, H. (2004). Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern. Lab. Anim.* **32**, 275–307.

Pfister, S., Steiner, K. A., Tam, P. P. L. (2007). Gene expression pattern and progression of embryogenesis in the immediate post-implantation period of mouse development. *Gene Expr. Patterns*, **7**, 558–573.

Piersma, A. H., Genschow, E., Verhoef, A., Spanjersberg, M. Q. I., Brown, N. A., Brady, M., Burns, A., Clemann, N., Seiler, A., and Spielmann, H. (2004). Validation of the postimplantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern. Lab. Anim.* **32**, 275–307.

Post, A., Crochemore, C., Uhr, M., Holsboer, F., and Behl, C. (2000). Differential induction of NF-kappaB activity and neural cell death by antidepressants in vitro. *Eur. J. Neurosci.* **12**, 4331–4337.

Pratten, M., Ahir, B. K., Smith-Hurst, H., Memon, S., Mutch, P., and Cumberland, P. (2012). Primary cell and micromass culture in assessing developmental toxicity. *Methods Mol. Biol.* **889**, 115–146.

Rajadhyaksha, A. M., Ra, S., Kishinevsky, S., Lee, A. S., Romanienko, P., DuBoff, M., Yang, C., Zupan, B., Byrne, M., Daruwalla, Z. R., Mark, W., Kosofsky, B. E., Toth, M., and Higgins, J. J. (2012). Behavioral characterization of cereblon forebrain-specific conditional null mice: a model for human non-syndromic intellectual disability. *Behav. Brain Res.* **226**, 428–434.

Reed, C. E., and Fenton, S. E. (2013). Exposure to diethylstilbestrol during sensitive life stages: a legacy of heritable health effects. *Birth Defects Res. C Embryo Today*. **99**, 134–146.

Reefhuis, J., Devine, O., Friedman, J. M., Louik, C., and Honein, M. A. ; National Birth Defects Prevention Study. (2015). Specific SSRIs and birth defects: Bayesian analysis to interpret new data in the context of previous reports. *BMJ*. **351**, h3190.

Rhinn, M., and Dolle, P. (2012). Retinoic acid signalling during development. *Development*, **139**, 843–858.



- Riebeling, C., Hayess, K., Peters, A. K., Steemans, M., Spielmann, H., Luch, A., and Seiler, A. E. (2012). Assaying embryotoxicity in the test tube: current limitations of the embryonic stem cell test (EST). challenging its applicability domain. *Crit. Rev. Toxicol.* **42**, 443–464.
- Roberts, S. S., Miller, R. K., Jones, J. K., Lindsay, K. L., Greene, M. F., Maddrey, W. C., Williams, I. T., Liu, J., and Spiegel, R. J. (2010). The Ribavirin Pregnancy Registry: Findings after 5 years of enrollment, 2003-2009. *Birth Defects Res. A Clin. Mol. Teratol.* **88**, 551–559.
- Ruiz-Villalba, A., Hoppler, S., and van den Hoff, M. J. (2016). Wnt signaling in the heart fields: Variations on a common theme. *Dev. Dyn.* **245**, 294-306.
- Russo, F. B., Jungmann, P., Cristina, P., and Beltrão, B. (2017). Zika infection and the development of neurological defects. 1–6.
- Rynn, L. *et al.* (2008). Update on overall prevalence of major defects - Atlanta, Georgia, 1978-2005. *Morb. Mortal. Wkly. Rep.*, **57**, 1–5.
- Sadar, M. D., Hussain, M., and Bruchovsky, N. (1999). Prostate cancer: molecular biology of early progression to androgen independence. *Endocr. Relat. Cancer.* **6**, 487–502.
- Sadler, T. W. (2012). Langman's Medical Embryology 12th edition. Taylor, C. (ed). Lippincott Williams & Wilkins, a Wolters Kluwer business, Philadelphia, PA.
- Sakai, Y., Meno, C., Fujii, H., Nishino, J., Shiratori, H., and Saijoh, Y. (2001). The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the antero-posterior axis within the mouse embryo. *Genes Dev.*, **15**, 213–225.
- Schardein, J. L., and Macina, O. T. (2006). Human Developmental Toxicants: Aspects of Toxicology and Chemistry. Boca Raton: CRC Press.
- Schaz, U., Föhr, K. J., Liebau, S., Fulda, S., Koelch, M., Fegert, J. M., Boeckers, T. M., and Ludolph, A. G. (2011). Dose-dependent modulation of apoptotic processes by fluoxetine in maturing neuronal cells: an in vitro study. *World J. Biol. Psychiatry* **12**, 89-98.
- Schmid, G., Guba, M., Ischenko, I., Papyan, A., Joka, M., Schrepfer, S., Bruns, C. J., Jauch, K. -W., Heeschen, C., and Graeb, C. (2007). The immunosuppressant FTY720 inhibits tumor angiogenesis via the sphingosine 1-phosphate receptor 1. *J. Cell. Biochem.* **101**, 259–270.

Schüler, L., Pastuszak, A., Sanseverino, T. V., Orioli, I. M., Brunoni, D., Ashton-Prolla, P., Silva da Costa, F., Giugliani, R., Couto, A. M., Brandao, S. B., and Koren, G. (1999). Pregnancy outcome after exposure to misoprostol in Brazil: a prospective, controlled study. *Reprod. Toxicol.* **13**, 147–151.

Scialli, A. R., Daston, G., Chen, C., Coder, P. S., Euling, S. Y., Foreman, J., Hoberman, A. M., and Hui, J. (2018). Rethinking developmental toxicity testing : Evolution or revolution ? Seiler, A., Visan, A., Buesen, R., Genschow, E., and 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.

Seiler, A. E., Buesen, R., Visan, A., and Spielmann, H. (2006). Use of murine embryonic stem cells in embryotoxicity assays: the embryonic stem cell test. *Methods Mol. Biol.* **329**, 371–395.

Seiler, A. E., and Spielmann, H. (2011). The validated embryonic stem cell test to predict embryotoxicity in vitro. *Nat. Protoc.* **6**, 961–978.

Sipes, N. S., Padilla, S., and Knudsen, T. B. (2011). Zebrafish: as an integrative model for twenty-first century toxicity testing. *Birth Defects Res. C Embryo Today* **93**, 256–267.

Skerjanc, I. S. (1999). Cardiac and skeletal muscle development in P19 embryonal carcinoma cells. *Trends Cardiovasc. Med.* **9**, 139–143.

Slaughter, S. R., Hearn-Stokes, R., van der Vlugt, T., and Joffe, H. V. (2014). FDA approval of doxylamine-pyridoxine therapy for use in pregnancy. *N. Engl. J. Med.* **370**, 1081–1083.

Sloot, W. N., Bowden, H. C., and Yih, T. D. (2009). In vitro and in vivo reproduction toxicology of 12 monoaminergic reuptake inhibitors: possible mechanisms of infrequent cardiovascular anomalies. *Reprod. Toxicol.* **28**, 270–282.

Sousa-Ferreira, L., Aveleira, C., Botelho, M., Álvaro, A. R., Pereira de Almeida, L., and Cavadas, C. (2014). Fluoxetine induces proliferation and inhibits differentiation of hypothalamic neuroprogenitor cells in vitro. *PLoS One* **9**, e88917.

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

Spielmann, H., and Liebsch, M. (2001). Lessons learned from validation of in vitro toxicity test: From failure to acceptance into regulatory practice. *Toxicol. Vitro*, **15**, 585–590.

- Starling, L. D., Sinha, A., Boyd, D., and Furck, A. (2012). Fetal warfarin syndrome. *BMJ Case Rep.* doi: 10.1136/bcr-2012-007344.
- Sun, B. K., Kim, J. H., Choi, J. S., Hwang, S. J., and Sung, J. H. (2015). Fluoxetine decreases the proliferation and adipogenic differentiation of human adipose-derived stem cells. *Int. J. Mol. Sci.* **16**, 16655-16668.
- Suzuki, N., Ando, S., Yamashita, N., Horie, N., and 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.
- Tada, M., and Heisenberg, C. P. (2012). Convergent extension: using collective cell migration and cell intercalation to shape embryos. *Development* **139**, 3897–3904.
- Tam, P. P. L., Loebel, D. A. F., and Tanaka, S. S. (2006). Building the mouse gastrula: signals, asymmetry and lineages. *Curr. Opin. Genet. Dev.*, **16**, 419–425.
- Tam, P. P., and Loebel, D. A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet.* **8**, 368–381.
- Tamashiro, D. A., Alarcón, V. B., and Marikawa, Y. (2008). Ectopic expression of mouse Sry interferes with Wnt/beta-catenin signaling in mouse embryonal carcinoma cell lines. *Biochim. Biophys. Acta.* **1780**, 1395-1402.
- Teter, C. J., Phan, K. L., Cameron, O. G., and Guthrie, S. K. (2005). Relative rectal bioavailability of fluoxetine in normal volunteers. *J. Clin. Psychopharmacol.* **25**, 74-78.
- Theunissen, P. T., and Piersma, A. H. (2012). Innovative approaches in the embryonic stem cell test (EST). *Front. Biosci.* **17**, 1965–1975.
- Tonk, E. C. M., Pennings, J. L. A., and Piersma, A. H. (2015). An adverse outcome pathway framework for neural tube and axial defects mediated by modulation of retinoic acid homeostasis. *Reprod. Toxicol.*, **55**, 104–113.
- Trosko, J. E., and Chang, C. -C. (2010). Factors to consider in the use of stem cells for pharmaceutical drug development and for chemical safety assessment. *Toxicology*, **270**, 18–34.
- Udechuku, A., Nguyen, T., Hill, R., and Szego, K. (2010). Antidepressants in pregnancy: a systematic review. *Aust. N. Z. J. Psychiatry.* **44**, 978-996.

- Ueno, N., and Greene, N. D. (2003). Planar cell polarity genes and neural tube closure. *Birth Defects Res. C Embryo Today* **69**, 318–324.
- Vargesson, N. (2009). Thalidomide-induced limb defects: resolving a 50-year-old puzzle. *Bioessays*. **31**, 1327–1336.
- Videla, S., Cebrecos, J., Lahjou, M., Wagner, F., Guibord, P., Xu, Z., Cabot, A., Encabo, M., Encina, G., Sicard, E., and Sans, A. (2013). Pharmacokinetic dose proportionality between two strengths (12.5 mg and 25 mg) of doxylamine hydrogen succinate film-coated tablets in fasting state: a single-dose, randomized, two-period crossover study in healthy volunteers. *Drugs. R. D.* **13**, 129–135.
- Waring, M. J., Arrowsmith, J., Leach, A. R., Leeson, P. D., Mandrell, S., Owen, R. M., Pairaudeau, G., Pennie, W. D., Pickett, S. D., *et al.* (2015). An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat. Rev. Drug Discov.* **14**, 475–486.
- Voronova, A., Coyne, E., Al Madhoun, A., Fair, J. V., Bosiljic, N., St-Louis, C., Li, G., Thurig, S., Wallace, V. A., Wiper-Bergeron, N., and Skerjanc, I. S. (2013). Hedgehog signaling regulates MyoD expression and activity. *J. Biol. Chem.* **288**, 4389–4404.
- Warkus, E. L., Yuen, A. A., Lau, C. G., and Marikawa, Y. (2016). Use of in vitro morphogenesis of mouse embryoid bodies to assess developmental toxicity of therapeutic drugs contraindicated in pregnancy. *Toxicol. Sci.* **149**, 15–30.
- Warkus, E. L., and Marikawa, Y. (2017). Exposure-based validation of an in vitro gastrulation model for developmental toxicity assays. *Toxicol. Sci.* **157**, 235–245.
- Weisberg, E., Winnier, G. E., Chen, X., Farnsworth, C. L., Hogan, B. L., and Whitman, M. (1998). A mouse homologue of FAST-1 transduces TGF beta superfamily signals and is expressed during early embryogenesis. *Mech. Dev.* **79**, 17–27.
- Wells, P. G., and Winn, L. M. (1996). Biochemical toxicology of chemical teratogenesis. *Crit. Rev. Biochem. Mol. Biol.* **31**, 1–40.
- Wemakor, A., Casson, K., Garne, E., Bakker, M., Addor, M. C., Arriola, L., Gatt, M., Khoshnood, B., Klungsoyr, K., Nelen, V., *et al.* (2015). Selective serotonin reuptake inhibitor antidepressant use in first trimester pregnancy and risk of specific congenital anomalies: a European register-based study. *Eur. J. Epidemiol.* **30**, 1187–1198.
- White, T. E. K., Bushdid, P. B., Ritter, S., Laffan, S. B., and Clark, R. L. (2006). Artesunate-induced depletion of embryonic erythroblasts precedes embryo lethality and teratogenicity in vivo. *Birth Defects Res. B Dev. Reprod. Toxicol.* **77**, 413–429.

White, T. E. K., and Clark, R. L. (2008). Sensitive periods for developmental toxicity of orally administered artesunate in the rat. *Birth Defects Res B Dev. Reprod. Toxicol.* **83**, 407–417.

Wilde, J. J., Petersen, J. R., and Niswander, L. (2014). Genetic, Epigenetic, and Environmental Contributions to Neural Tube Closure.

Willet, M. N., Hayes, D. K., Zaha, R. L., and Fuddy, L. J. (2012). Social-emotional support, life satisfaction, and mental health on reproductive age women's health utilization, US, 2009. *Matern. Child Health J.* **16**, 203-12.

Wise, L. D. (2016). Numeric estimates of teratogenic severity from embryo-fetal developmental toxicity studies. *Birth Defects Res. B Dev. Reprod. Toxicol.* **107**, 60–70.

Wobus, a M. 2001. "Potential of Embryonic Stem Cells. " *Molecular Aspects of Medicine* 22 (3): 149–64., Wobus, Anna M., and Peter Löser. 2011. "Present State and Future Perspectives of Using Pluripotent Stem Cells in Toxicology Research. " *Archives of Toxicology* 85 (2): 79–117. doi:10. 1007/s00204-010-0641-6., Wobus, A. M. *et al.* (1988). Specific effects of nerve growth factor on the differentiation pattern of mouse embryonic stem cells in vitro. *Biomed. Biochim. Acta*, **47**, 965–973.

Wong, D. T., Bymaster, F. P., and Engleman, E. A. (1995). Prozac (fluoxetine, Lilly 110140), the first selective serotonin uptake inhibitor and an antidepressant drug: twenty years since its first publication. *Life Sci.* **57**, 411-441.

Wong, D. T., Reid, L. R., and Threlkeld, P. G. (1988). Suppression of food intake in rats by fluoxetine: comparison of enantiomers and effects of serotonin antagonists. *Pharmacol. Biochem. Behav.* **31**, 475-479.

Xing, J., Toh, Y. C., Xu, S., and Yu, H. (2015). A method for human teratogen detection by geometrically confined cell differentiation and migration. *Sci Rep.* **5**, 10038.

Yamaguchi, T. P., and Rossant, J. (1995). Fibroblast growth factors in mammalian development. *Curr. Opin. Genet. Dev.* **5**, 485–491.

Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hübner, K., and Schöler, H. R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881–894.

Yu, R., Miyamura, N., Okamoto-Uchida, Y., Arima, N., Ishigami-Yuasa, M., Kagechika, H., Nishina, H. (2015). A modified murine embryonic stem cell test for evaluating the teratogenic effects of drugs on early embryogenesis. *PLoS One.* **10**, e0145286.

Yu, J., and Thomson, J. A. (2008). Pluripotent stem cell lines. *Genes. Dev.* **22**, 1987–1997.

Zhao, J., Krafft, N., Terlouw, G. D., and Bechter, R. (1993). A model combining the whole embryo culture with human liver S-9 fraction for human teratogenic prediction. *Toxicol. In Vitro* **7**, 827–831.