Endless Possibilities: Stem Cells and the Vision for Toxicology Testing in the 21st Century

Robert E. Chapin* and Donald B. Stedman

Developmental and Reproductive Toxicology, Drug Safety R&D, Pfizer, Inc., Groton, Connecticut 06365

Received July 31, 2009; accepted August 18, 2009

The National Research Council’s (NRC) toxicity testing vision lays out a bold future for our field. It depends heavily on computational algorithms based on the latest knowledge of cellular biochemistry and protein interaction pathways, exposing human cells to novel compounds in vitro, and being able to understand the changes seen. At the same time, significant strides are being made in our understanding of the control, production, and “behavior” of stem cells. While stem cells offer seemingly limitless possibilities for regenerative medicine, they have already delivered new assays to predict embryo-fetal developmental toxicity in vitro. In addition to providing a model of cells undergoing differentiation and proliferation, stem cells will play a major role by giving rise to many of the differentiated cell types on which this new vision depends. These will not be pure populations of single cell types but mixtures of cells much more representative of tissues in vitro. Moving from cells alone in a culture dish toward the more physiological condition of multiple cell types being able to interact to maintain homeostasis in the face of a disequilibrating force (like a toxic exposure) will lead us toward more useful and correct predictions of in vivo toxicities. Despite the seemingly insurmountable hurdles, persistence and creativity are on our side. We expect that a long series of successive iterations of predictive models will eventually yield a working process that approximates the NRC’s vision and delivers on the promise of faster evaluation of chemicals with reduced animal use.

Key Words: embryonic stem cells; induced pluripotent stem cells; predictive toxicology; toxicity testing.

The vision articulated in “Toxicity Testing in the 21st Century” leans heavily on cell-based assays, knowledge of systems biology and the pathways of cell response to toxicants, and to a large extent on computational agility. We agree that this is an excellent vision, toxicology’s own version of the Grand Unification hypothesis in physics, a seeing of things that are not and saying “Why the heck not?!” At this point in time, the main value of this vision is as a framework providing the outlines of the way toxicity testing would ideally be done years from now. Like all visionary frameworks, it has serious and major gaps, but we understand and applaud the benefits of such a vision. These gaping holes serve to identify where we collectively should be directing our efforts. Gap filling can unite and guide the efforts of everyone in a field because as soon as concrete advances are made in each area, they can be put to immediate use and deliver immediate improvement to the model.

This installment in the series of Forum articles deals with one of the testing tools articulated in the vision and in the initial discussion piece by Andersen and Krewski (2009): stem cells. Embryonic stem cells are those cells which come from the inner cell mass of the blastocyst and from which the entire body arises. They can divide nearly endlessly and differentiate into any cell type. Stem cells are the new darlings of both basic science and therapeutics, and interest in them has been exploding (Fig. 1). They are the yet-to-be-elected but popular politician, brimming with promise and hope: they are touted as providing new cures for previously untreatable diseases, they might replace worn-out cells and provide new therapies for aging, they can theoretically differentiate into any final cell type (once we learn the cues required to guide them) and thus provide new substrate for developing new medicines . . . the possibilities seem limitless. As with the politician, there can be a significant gulf between promise and delivery, but the promises are so beguiling, so “alluring,” that we elect them in the hopes that miracles still occur.

The recent production of induced pluripotent stem (iPS) cells adds another layer of luster to the appeal (Maherali and Hochedlinger 2008; Takahashi and Yamanaka, 2006; Zhao and Daley, 2008). These cells can come from a (or theoretically any) terminally differentiated cell in the body, and by transduction with a few transcription factors (TFs) (Oct4, Nanog, Klf4, and Sox2), they can be returned to their undifferentiated pluripotent state. The big appeal of these cells...
is (1) they are derived from an adult cell and do not require the destruction of an embryo, (2) because they could in theory be made from anyone’s cells, they could avoid immune rejection if the starting cells were taken from the intended ultimate recipient, and (3) they could allow the creation of cell lines from individuals with diverse genetic backgrounds for study of specific diseases. If iPS cells prove to be as robust as human embryonic stem cell in generating differentiated cells, they could open up new doors to the understanding of the genetic versus environmental factors that contribute to an individual’s response to a drug or chemical. This method has implications so profound that the journal *Science* proclaimed “Reprogramming” (the derivation of iPS cells from differentiated cells) the Breakthrough of the Year last year (Vogel, 2008).

As we think about how stem cells might contribute to the tox testing vision, there are two additional items to remember: (1) Nature abhors 2D monocultures, both macroscopically and microscopically; there is no such thing in vertebrates as a pure population of a single kind of cell connected to other cells only in two dimensions. Thus, it is not surprising that any given population of stem cells, even highly purified embryonic stem cells, is heterogeneous, with varying amounts of the pluripotency-inducing TFs in different cells, which in turn leads to different behaviors of those cells in the cultures and *in vivo* (Graf and Stadtfeld, 2008). And (2) most adult tissues contain stem cells specific for that tissue type. Some of these turn over much more slowly or undetectably (brain and heart) (e.g., Hoffman, 2008; Hoogendoorn et al., 2008), while others turn over quickly and are responsible for the high rates of cell turnover (i.e., testis, skin, and gut), which in turn leads to different behaviors of those cells in the cultures and *in vivo*. (Graf and Stadtfeld, 2008). These tissue-specific stem cells are also heterogeneous, existing in multiple states characterized by different degrees of expression of various TFs (Graf and Stadtfeld, 2008). In this review, we are mostly referring to pluripotent embryonic stem cells or iPS cells.

As with any promising new technique, there is a combination of irrational exuberance and some truth to all the rumors. Guiding stem cell differentiation to a desired final differentiated cell type is a combination of skill and luck. The goal of the field is still to generate final differentiated cells from pluripotent stem cells in culture because only cultures can be scaled to meet the volume needs of the testing which is envisioned, and this is an opportunity to free ourselves from the costs and logistical and moral constraints around using animals. Thus far, success in producing an enriched population of a specific cell type requires a specific cocktail of extrinsic growth factors or transfection with specific TFs and a specific external environment, either feeder layers or a particular 3D culture condition or implantation into an injured *in vivo* model, to stimulate the final differentiation (reviewed in Godier et al., 2008) (Table 1). We are still some ways from having a straightforward protocol to generate pure adult hepatocytes, for example (Banas et al., 2006; Soto-Gutierrez et al., 2008). It is possible to make cells that have many characteristics of cells along the lineage of hepatocytes, but a protocol that produces adult terminally differentiated hepatocytes is still future tense. Similar challenges currently exist in making insulin-producing cells (reviewed in Raikwar and Zavazava, 2008). Stem cell–derived cardiomyocytes are another example of a differentiated cell type that has a number of cardiac-specific characteristics (protein markers, ion channels, pacemaker, atrial, and ventricular action potentials). However, these cardiomyocytes are still more fetal like than adult. There has been more success with other cell types: pure populations of

---

**TABLE 1**

Examples of cells derived from embryonic stem cells

<table>
<thead>
<tr>
<th>Cell types</th>
<th>Phenotype</th>
<th>Markers</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural</td>
<td>Dopaminergic</td>
<td>Tyrosine hydroxylase, serotonin</td>
<td>Lee et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>Serotonergic</td>
<td>Serotonin</td>
<td>Kim et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>HB9, Isl1</td>
<td>Wichterle et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>GABA neurons</td>
<td>FOXG1B</td>
<td>Barbieri et al. (2003)</td>
</tr>
<tr>
<td>Astrocytes</td>
<td>GFAP</td>
<td>Barbieri et al. (2003)</td>
<td></td>
</tr>
<tr>
<td>Glial cells</td>
<td>GFAP</td>
<td>Barbieri et al. (2003)</td>
<td></td>
</tr>
<tr>
<td>Cardiomyocytes</td>
<td>Cardiac cells</td>
<td>MHC and Nkx2.5</td>
<td>*Wobus et al. (2002)</td>
</tr>
<tr>
<td>Hepatic</td>
<td>Hepatocytes</td>
<td>ALB, AAT, and ATP</td>
<td>Hamazaki et al. (2001)</td>
</tr>
<tr>
<td>Vascular</td>
<td>Endothelial</td>
<td>PECAM (CD31), Flt-1, VE-cadherin, and CD34</td>
<td>Vittet et al. (1996)</td>
</tr>
<tr>
<td>Islet</td>
<td>β-cell</td>
<td>Insulin I, insulin II, islet amyloid, and GLUT-2</td>
<td>Lumelsky et al. (2001)</td>
</tr>
<tr>
<td>Skeleton</td>
<td>Chondrocyte</td>
<td>Type II collagen</td>
<td>Kramer et al. (2000)</td>
</tr>
<tr>
<td>Osteoblast</td>
<td>Osteoblast</td>
<td>Mineralized matrix</td>
<td>Bielby et al. (2004)</td>
</tr>
</tbody>
</table>

*Note.* This table presents some of the cells that have been produced from stem cells *in vitro* and provides the markers used to confirm their presence in the cultures. GABA, gamma-aminobutyric-acid; GFAP, glial fibrillary acidic protein.
functional dopaminergic neurons have been produced in vitro, for example (Kim et al., 2002).

There are a number of efforts worldwide which are currently focused on the application of stem cells to disease treatment and safety testing, either as stem cells or as their differentiated daughters. The Stem Cells for Safer Medicines consortium in the United Kingdom is a public-private partnership aimed at converting stem cells into hepatocytes and generating predictive safety screens. Geron (Menlo Park, CA), Cellartis (Göteborg, Sweden), and VistaGen Therapeutics (South San Francisco, CA) are companies which each currently provide or are developing safety screens for in vivo toxicities. In addition, Geron is pursuing the production of osteoblasts for treating osteoporosis, chondrocytes for treating osteoarthritis, and hepatocytes for multiple therapeutic indications. Cellular Dynamics (Madison, WI) is another company developing a cardiomyocyte-based assay for cardiac safety prediction. In the regrettable current vernacular: “This is a thriving private sector space.”

One issue of concern is which stem cells will be most useful for fitting into this vision. The answer is it is too soon to tell. Again, theoretically, any truly pluripotent stem cell should be able to fill in and provide any differentiated cell type. However, the practical matter of not yet knowing all the necessary cues to produce the desired cell types will likely lead us to employ tissue-specific stem cells or mixes thereof as a way to shortcut the occult early differentiation process. This will be an area of much trial and error and will be a secure home for visionary pragmatism for the foreseeable future.

The challenges of producing “pure” populations of a final cell type are not surprising. A pure population may not even be what is needed. Recall that cells in vivo do not develop or live in pure monoculture but they live and function admixed with other cell types as a tissue, and we predict that this heterogeneity will eventually be recognized as crucial to the sustained differentiated function of any cell type. Thus, flasks of pure hepatocytes, for example, is the wrong vision. A better goal might be mixed 3D cultures of hepatocytes, Kupfer cells, bile ducts, and fibroblasts because only a mixed population in three dimensions will best mimic the in vivo responses we seek. We might call this a tissue “doppel.” Liver spheroids are one step toward this goal and have been shown to better replicate the in vivo response of the liver to many exposures, nanoparticles being only one recent example (Lee et al., 2009).

One other useful characteristic of stem cells is that iPS cells generated from humans with specific diseases maintain some of the programming characteristic of that disease (Beqqali et al., 2009; Ebert et al., 2009). This implies that we can (and should) eventually obtain or generate iPS cells from a wide variety of people, encompassing the broad spectrum of metabolic abilities, drug susceptibilities, resistance or susceptibility to disease, etc. Conceptually, this could yield enormous benefits. As with all conceptual gifts, however, the details of harnessing this diversity and squeezing something useful from it, even a fraction of the promise, seems impressively complex. George Church’s Personal Genome Project at Harvard is doing exactly this and will help highlight the challenges and their solutions.

So how do stem cells fit into the vision? In two ways: (1) by being differentiated into cultures of different “human” cell types whose response to putative toxicants can then be assessed and (2) by being evaluated for their response to toxicants in their undifferentiated state (or during differentiation) during toxicant exposure. The specific allure is that we can use human cells and thus avoid the extrapolation of responses from animals, and we can do this with a theoretically limitless supply of cells. The goal might be a production process which generates cultures of differentiated cells representative of various adult tissues which would be exposed to the chemical of concern and the responses measured. This goal assumes continued progress in understanding the basic biology of differentiation for each tissue (which seems reasonable) in a time frame short enough to be usable (perhaps more of a stretch), and it assumes that we know the limitations of each system. These limitations will become more and more apparent as we push these cultures to meet our expectations, which are probably well beyond what the cells can reasonably do. So while stem cells themselves are part of the vision, arguably their greater contribution is to produce the differentiated cells that we want and need.

What will it take to realize the promise inherent in these cells?

1) Robust cell lines which are prolific, easy to maintain, differentiate appropriately and readily, which recapitulate in vitro the known in vivo responses to well-known stimuli (both trophic and toxic), and which yield results which are reproducible across replicates and laboratories.

2) A known method for the guided differentiation into, say, 15 desired “tissue doppels,” which will include the known induction cocktail and environmental requirements for producing adult-state tissue recapitations in vitro.

3) To account for differences in susceptibility, there needs to be a well-thought-through process for obtaining cells from a spectrum of genotypes and then applying these in a thoughtful and systematic way to the predictive process. Again, this should be achievable.

4) An appropriately scaled production and distribution system for all these different types of cells. This is “merely” a logistical challenge, not necessarily requiring scientific breakthroughs to achieve, although the logistics of this are daunting.

It is worth noting that all of the above use stem cells as a precursor factory, and once the cells start to differentiate, they are no longer stem cells. In this vision, the guided differentiation of stem cells, and their production at scale, becomes an “enabling technology.”

Of these five requirements above, none is impossible. None of them exists yet, either, but the record of human innovation
and Ray Kurzweil’s strong documentation of the logarithmic speed of advance in every area (Fig. 2, and Kurzweil, 2005) give us great confidence that these requirements are achievable within the next 5–8 years.

Now, in contrast to using them as a factory to make other cells, can they be exploited for what they naturally and spontaneously want to do? The answer is “Yes.” What do stem cells do? They divide and differentiate. This is what normally happens in an embryo, so it is hardly surprising, then, that stem cells are being used “as stem cells” to help predict developmental toxicity. We would not reprise the whole process by which the European Commission for the Validation of Alternative Methods evaluated and “validated” the embryonic stem cell test because that is well documented already (Buesen et al., 2009; Genschow et al., 2004; Paquette et al., 2008; Peters et al., 2008; Seiler et al., 2004). It is enough to know that stem cells are already being exploited for doing what comes naturally. It is the possibility of using them to produce cells and cell culture models that do not yet exist which is so alluring.

How does one measure responses in the stem cells themselves? One way is to do what the tox testing vision suggests and focus on pathways and networks of genes, with the pattern of response relating to an in vivo toxicity profile. This is just now being explored in stem cells. Figure 3 is an example of such efforts. In our own laboratory, we have started to evaluate the effects of teratogens and non-teratogens on the 17 signaling pathways identified by a National Academy sub-committee in 2000 (NRC, 2000). This is simply one way of showing gene ontologies, grouped by cellular function. Assessing the changes due to an exposure and then assembling those data into “knowledge” is when the mental plow must be set deep and (for the time being) progress is slow. This step of synthesis and assembly is another opportunity for improvement, as it will be required of all network and pathway analyses.

It is clear that stem cells can play a significant role in the realization of the tox testing vision. Since we have some
experience in working with in vitro systems and using them to predict toxicity, we would like to turn now to some considerations of the vision itself.

This tox testing vision invokes the power of pathways and asserts that knowing which pathways are disrupted will help separate toxic compounds from nontoxic ones. Being long-time students of scientific progress, we are certain that reality will be much more complicated than this. There is still much to learn about how pathways relate to phenotype and a toxicity seen in vivo. And even leaving all that aside, there is much more to learn about how physiology works and how organs and tissues interact to maintain biochemical stability in the gale of exogenous chemicals. Even with all we know, we realize that our ignorance is profound and deep. On the other hand, it may be true that in the final version of this testing scenario, we would not need to know how a toxicity will manifest but would only need to know which tissue doppels in vitro are sufficiently affected to pass over the threshold of change into toxicity. Perhaps, we will not need to reconstruct all the steps leading from reduced neuronal steroid sensitivity to increased ovarian steroid output to altered estrous cycle (persistent estrus) to infertility; eventually, it may be that seeing the neuronal change will be enough to flag a compound as potentially toxic and lead to its testing in animals. Indeed, in a perverse way, this testing system might actually help short-circuit many future mechanism-of-action studies because the animal tests would be triggered by an effect noted in a certain tissue doppel in vitro, which may or may not be the tissue that manifests the toxicity in vivo. Assuming few false positives (always a dangerous assumption), we could see this leading to a small revolution in our understanding of how cells and tissues link together physiologically in vivo. So there is much to learn about how to relate pathways to phenotype in vivo, even while we grant that this may not be critical to implementing the pathways-based testing vision.

In addition, the body does automatically what a testing structure would have to do consciously: integrate a number of physiological processes in series. Every in vitro assay is an approximation of what happens in vivo, and it is never 100% predictive. Modeling the numerous events produced by an exogenous compound in vivo could build error upon error. For example, to turn a pathways-based hazard identification into a risk assessment for free-range humans, we will likely be modeling things like uptake across an epithelial barrier, metabolism, excretion, and an organ-specific toxicity response. For the sake of argument, let us simplify and model only those four processes. If each prediction has an 85% chance of being correct (and in our experience, this is greater than the performance of most predictive assays), then we have a $(0.85)^4 = 52\%$ chance of running something through all four models in series and correctly predicting what will be seen in vivo. This seems hardly worth all the effort to set up this massive process. We should not just amplify the weaknesses of using multiple assays in series.

All the foregoing is why we caution against overoptimism. We should not confuse perspicacity with propinquity. That is, just because we can “see” an object (through the powerful telescope of the tox testing vision) does not mean that this object is “close.” “Seeing it” is different from “being there,” and there are many hurdles to cross before the vision is realized.

Nonetheless, we remain strong supporters of such a vision. It provides a guide and a framework that allows the immediate application of any advance in understanding or insight into a mechanism of toxicity or even basic biology and thus helps drive all science forward. It should (eventually) help reduce animal use while generating more data, which are two noble goals. And it is a deeply compelling vision, rooted in the assumption that more knowledge, rationally assembled and compassionately used, can improve the human condition. It is a challenge worthy of the greatest minds among us.

REFERENCES


