An Evaluation of new Bioinformatics Techniques to replace the Teratoma pluripotency assay

David Elan Feldman

5/6/16

SPECIFIC AIMS

Recent advances in deriving human induced pluripotent stem cells (iPSCs) from somatic cells hold great promise in the field of regenerative medicine. iPSCs are, by nature, not subject to many of the ethical and legal dilemmas associated with embryonic stem (ES) cells and hence provide a means for the scientific community to clear what had previously been a major research bottleneck. Notwithstanding, as iPSC research has continued to expand and develop, difficulties surrounding iPSC research have arisen in their own right. As the number of iPSC cell lines that are currently being researched increases and more and more cell lines become catalogued in cell banks and repositories, establishing the authenticity, functional symmetry, and definitive pluripotency of individually researched iPSCs will become increasingly difficult and important.

Though teratoma formation in immunosuppressed mice is largely considered to be the "gold standard" for assessing pluripotency, little consistency exists in research protocols for conducting teratoma tests despite efforts to establish such protocols. Moreover, conducting teratoma tests presents ethical concerns, is time intensive, and can be expensive. Using teratoma formation to screen for pluripotency is not ideal because of the reasons stated above; resent research has focused on developing alternative assays to confirm iPSC pluripotency. Researchers have suggested that viable alternatives include Embryoid Body (EB) assays, chorioallantoic membrane chicken egg assays, and computer based bioinformatics predictive modeling assays.

PluriTest, an open source bioinformatic package designed to assess the functional pluripotency of ES cells and iPSCs based upon genomic signatures, was developed in 2011 but has failed to gain substantial traction. In 2013 Life Sciences (now Thermo Fischer Scientific) released the TaqMan® hPSC Scorecard, and pluripotency assessment product including a protocol and bioinformatics assay software package. Though this assay failed to gain traction within the research community because of it's difficulty of use and perceived inaccuracy, in late 2015 researchers developed a new protocol and software suite that claims to eliminate problems entirely. Notwithstanding, PluriTest and the TaqMan® hPSC Scorecard assay have yet to be independently compared and evaluated for accuracy.

The major objective of this proposal is to conduct a thorough analysis of bioinformatics based pluripotency assays to definitively establish a new standard for iPSC and ES cell pluripotency assessment. *If successful, this research would advance iPSC research by creating a standardized platform for pluripotency assessment and by decreasing the costs and time necessary to assess pluripotency.*

AIM 1: Comparatively evaluate the results obtained using the PluriTest bioinformatic assay protocol, the revised TaqMan® hPSC Scorecard protocol and the standardized (Gropp et al. 2012) Teratoma tumor formation protocol using 200 cell lines including some of the most widely circulated iPSC lines, ES cell lines, and non pluripotent stem cell lines. We will independently assess he results of the computer based models by comparing the outputs with teratoma assay results. Though all three assay protocols leave room for subjective assessment in their final analysis of pluripotnecy, we will largely seek to determine if different assays may lead researchers to different conclusions so far as whether a cell line is pluripotent, not pluripotent, or tests inconclusively.

AIM 2: Independently assess the costs, laboratory man hours, and experimental time required for each of the aforementioned pluripotency assays. In doing so, we seek to weigh all relevant factors necessary to establish a next-generation pluripotency assay standard. We we track the aggregate hours worked in lab for each assay method, the average time spent from experimental initiation to completion for each assay method, and track the costs of running each assay.

RESEARCH STRATEGY

(a) Significance

Several fundamental procedural issues currently plague induced pluripotent stem cell (iPSC) research (1,13). Though the general consensus is that methods such as Short Tandem Repeat (STR) profiling must be used to accurately assess cell line authenticity, there is no general consensus on specific protocols for doing so and, more importantly, no indication for how frequently such profiling should be done to ensure authenticity (4,10,13). What is more, STR profiling alone cannot guarantee the functional symmetry of cell lines because of additional factors such as epigenetic changes, metabolic reprogramming, and mitochondrial alterations that may lead to cell line instability. Though teratoma formation in immunosuppressed mice is largely considered to be the "gold standard" for assessing pluripotency, little consistency exists in research protocols for conducting teratoma tests despite efforts to establish such protocols (1,3,5,12). Moreover, conducting teratoma tests presents ethical concerns, is time intensive and can be expensive (1).

The focus of the research proposed herein is to comparatively and independently evaluate the two most prominent computer based predictive modeling pluripotency assays available for iPSC researchers: PluriTest and the TaqMan® hPSC Scorecard (6,11). We hope to establish a clear path forward in establishing a non teratoma-based protocol as a next generation "gold standard" for pluripotency assays. In doing so we will aid in the development of iPSC regenerative therapy research as a whole through eliminating inconsistencies and promoting the usage of beneficial technologies. If we are able to definitively conclude that a bioinformatics based assay produces results that are of sufficient quality and our cost and time analysis plays out as we expect it to, we hope that our results will persuade the research community to utilize next generation bioinformatic methods in lieu of teratoma assays and save what amounts to an aggregate sum of millions of dollars and many years (1).

(b) Research Plan – Approach

b.1 Replacement of the teratoma assay. The experimental generation of teratomas in immunosuppressed mice, as a means of testing for the pluripotency of stem cells, has occurred for decades and is, to this day, largely recognized as "gold standard" for pluripotency assessment (1,3). Because of the increasing research interest in iPSCs, and the pressing need to screen for pluripotency in iPSC development, being able to accurately and easily screen for cell pluripotency has never been more important.

The teratoma assay is practically and financially limiting to current iPSC research and is considered by some to be ethically objectionable (1). Though the teratoma assay is often used to test for pluripotency in human cell lines, some researchers have recently questioned its efficacy as a valid assay because it consists of exposing cells to an exogenous *in vivo* environment. What is more, because of the many experimental factors involved with conducting a teratoma assay, which include vermin living environment, means of injection and number of injected cells, and actual evaluation of any teratoma formed, experimental implementation of the teratoma assay has been remarkably inconsistent across published research.

Hence, the overarching purpose of the research proposed herein it to promote a viable path forward (and away from the teratoma assay) for pluripotency assays in stem cell research. We intend to do so through independently and comparatively assessing the validity of new bioinformatics based genetic assays that used computer predictive modeling to screen for pluripotency.

b.2 The PluriTest assay. The PluriTest assay, developed in 2011, was constructed using machine learning algorithms, run on a database of genome wide transcriptional profiles of ~450 samples including diverse stem cell preparations from multiple laboratories, differentiated cell types, and developing and adult human tissues, to predict the presence or absence of pluripotent features in an unknown cell sample (6). Given a sequenced transcriptome, the makers of PluriTest claim to be able to identify the pluripotency of a cell line with 98% sensitivity and 100% specificity.

Notwithstanding, PluriTest has not been widely adopted by stem cell researchers as an acceptable stand-alone assay; though PluriTest has been used on more than 13,000 data sets since its publication journals to this day often require teratoma testing to definitively prove cell line pluripotency (6).

In our research, we seek to definitively and independently establish the accuracy of PluriTest as a pluripotency assay. Because of the lack of a specific protocol for end to end assessment of pluripotency using PluriTest, we seek to develop and publish such a protocol. Should our results indicate PluriTest to be an assay that is as highly sensitive and specific as it's makers claim, we hope to eliminate the need for auxiliary teratoma tests (7,9).

b.3 The TaqMan hPSC ScoreCard assay. Like the PluriTest assay, the ScoreCard approach evaluates the molecular signature of pluripotency and expression signatures that indicate functional pluripotency, defined as differentiation into each of the three germ layers (11). Unlike the PluriTest assay, which provides a final step analytics solution of an entire transcriptome, the TaqMan ScoreCard assay is sold as a kit for end to end pluripotent assessment and evaluates a specific set of genes via qPCR measurement. The TaqMan hPSC ScoreCard is likely substantially more expensive than the PluriTest assay; a kit to run four samples costs in excess of $400 and requires proprietary equipment. However, the TaqMan hPSC ScoreCard, as opposed to presenting a direct measure of "pluripotency" as an output, produces an array of results that is then manually compared to baseline standards in order to make a pluriponecy evaluation. Though this may make the ScoreCard approach more difficult to initially use, it almost certainly make it more adaptable to changing ideology on the importance of various genetic markers in establishing pluripotency (11). Nevertheless, the methodlology used to develop the TaqMan hPSC ScoreCard assay is not publicly available, where as little independent assessment or evaluation of the 2013 assay has been performed, virtually no assessment of the updated 2015 protocol has been preformed.

In our research, we seek to definitively and independently establish the accuracy and reliability of the TaqMan hPSC ScoreCard assay in assessing the pluripotency of stem cells. Should our results indicate the the TaqMan hPSC ScoreCard assay to be an assay that is as accurate and informative as to the pluripotency of a cell line as it's makers claim, we hope to eliminate the need for auxiliary teratoma tests within future research using the ScoreCard pluripotency approach, promoting regenerative therapy research as a whole (8-9). Additionally, we wish to assess the long term financial and practical sensibility of using the ScoreCard approach over using the traditional teratoma approach.

(c) Experimental Design and Methods

AIM 1: Comparatively evaluate the results obtained using the PluriTest bioinformatic assay, the revised TaqMan® hPSC Scorecard protocol and the standardized (Gropp et al. 2012) Teratoma tumor formation protocol using 200 cell lines including some of the most widely circulated iPSC lines, ES cell lines, and non pluripotent stem cell lines (3).

Rationale

As indicated above, the primary rationale behind conducting this analysis is that creators of bioinformatics based pluripotency assays indicate that they may be capable of beneficially replacing teratoma pluripotency assays, but have yet to be independently evaluated and have yet to be widely accepted. Because of the greatly improved efficiency and financial efficiency prospects that computer-based pluripotency assays offer over traditional teratoma assays, and because of increasingly high need for pluripotency assays due to iPSC research, detailed investigation of currently available assays is warranted.

Experimental Design and Methods

In order to begin conducting this experiment we would determine a set of 200 readily available stem cell lines that we would acquire to provide an adequate breadth of cell lineage for our experimentation. It will be essential to acquire multiple ES cell samples, iPSC cells of many different origins, and developing and adult tissue stem cells. Once an adequate collection has been acquired and developed to cell counts adequate for maintenance, each cell line would be tested for pluripotency using the three different methodologies listed above. To highlight key details of each assay protocol: i) a standardized PCR method for the collection of complete cell line transcriptomes for the PluriTest assay would be developed and published ii) The revised TaqMan® hPSC ScoreCard assay will be followed exactly as published and iii) the Gropp et. al teratoma assay would be followed exactly as published (3,7,11).

In order to comparatively evaluate results, the outputs of each assay protocol would be converted to a 0 to 20-point scale, with 0 indicating the highest degree of cells that test as definitively not pluripotent and 20 indicating the highest degree of cells that test as definitively pluripotent. After each cell line has been scored using each assay, a statistical analysis would be conducted to determine each bioinformatics test's deviation from the current "gold standard" results.

Expected Results – Benchmarks for Sucess

We expect our results to definitively indicate one of the following (both of which are beneficial to the progression of pluripotency testing methods and iPSC research in general):

1) That either one or both of the bioinformatics based pluripotency assays currently available are adequately accurate so as to replace the teratoma assay as the "gold standard" for pluripotency assessment.
2) That neither assay will be capable of replacing the teratoma pluripotency assay. This would indicate that other assays currently under development must be developed to fruition.

Potential Problems and Alternative Strategies

We may run into problems acquiring a collection so large as the stated 200 stem cell lines. In this case we may consider lowering the number of cell lines experimented on pursuing pre existing data from other labs to be included in our analysis. Tests may (and may frequently) fail entirely during our experimental process in which case repeats should be conducted. Additionally, our experimentation may face financial constraints that could limit the number of samples we could test.

AIM 2: Independently assess the costs, laboratory man hours, and experimental time required for each of the aforementioned pluripotency assays.

Rationale

Because we have reason to believe that both the PluriTest assay and the TaqMan® hPSC ScoreCard will be able to adequately assess the pluripotency of a stem cell line, we look to develop other factors to differentiate the tests. Specifically, we would like to compare the time spent in lab conducting each test and the dollars spent conducting each test.

Experimental Design and Methods

As the experiment described in Aim 1 is being conducted, we will maliciously document material usage time spent in lab, and total time needed for each assay. Repeats due to experimental failure will also be weighed into the result. After all information is collected, it will be analyzed and processed so as to convey a useful and quantitative result.

Expected Results –Benchmarks for Sucess

1) We expect that traditional teratoma testing will be the most time, labor, and capital intensive.
2) We expect that PluriTest will be the cheapest test to perform but will require more man hours and will take a longer time that the TaqMan® hPSC ScoreCard test.

Potential Problems and Alternative Strategies.

Certain elements of our analysis may be difficult to quantify. For example, it is likely that the time and money necessary to raise the mice for the teratoma assay may be difficult to quantify or distinguish between if the mice are not raised in-house.

Works Cited

1. Buta, C., David, R., Dressel, R., Emgård, M., Fuchs, C., Gross, U., Healy, L., Hescheler, J., Kolar, R., Martin, U., Mikkers, H., Müller, F. J., Schneider, R. K., Seiler, A. E., Spielmann, H., and Weitzer, G. (2013) Reconsidering pluripotency tests: do we still need teratoma assays?, *Stem Cell Res 11*, 552-562.
2. De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H. G., Lensch, M. W., Lujan, E., Pei, D., Rossant, J., Wernig, M., Park, P. J., and Daley, G. Q. (2015) Hallmarks of pluripotency, *Nature 525*, 469-478.
3. Gropp, M., Shilo, V., Vainer, G., Gov, M., Gil, Y., Khaner, H., Matzrafi, L., Idelson, M., Kopolovic, J., Zak, N. B., and Reubinoff, B. E. (2012) Standardization of the teratoma assay for analysis of pluripotency of human ES cells and biosafety of their differentiated progeny, *PLoS One 7*, e45532.
4. Lenz, M., Goetzke, R., Schenk, A., Schubert, C., Veeck, J., Hemeda, H., Koschmieder, S., Zenke, M., Schuppert, A., and Wagner, W. (2015) Epigenetic biomarker to support classification into pluripotent and non-pluripotent cells, *Sci Rep 5*, 8973.
5. Müller, F. J., Goldmann, J., Löser, P., and Loring, J. F. (2010) A call to standardize teratoma assays used to define human pluripotent cell lines, *Cell Stem Cell 6*, 412-414.
6. Müller, F. J., Schuldt, B. M., Williams, R., Mason, D., Altun, G., Papapetrou, E. P., Danner, S., Goldmann, J. E., Herbst, A., Schmidt, N. O., Aldenhoff, J. B., Laurent, L. C., and Loring, J. F. (2011) A bioinformatic assay for pluripotency in human cells, *Nat Methods 8*, 315-317.
7. Nestor, M. W., and Noggle, S. A. (2013) Standardization of human stem cell pluripotency using bioinformatics, *Stem Cell Res Ther 4*, 37.
8. Phillips, M. D., Kuznetsov, S. A., Cherman, N., Park, K., Chen, K. G., McClendon, B. N., Hamilton, R. S., McKay, R. D., Chenoweth, J. G., Mallon, B. S., and Robey, P. G. (2014) Directed differentiation of human induced pluripotent stem cells toward bone and cartilage: in vitro versus in vivo assays, *Stem Cells Transl Med 3*, 867-878.
9. Robinton, D. A., and Daley, G. Q. (2012) The promise of induced pluripotent stem cells in research and therapy, *Nature 481*, 295-305.
10. Schlaeger, T. M., Daheron, L., Brickler, T. R., Entwisle, S., Chan, K., Cianci, A., DeVine, A., Ettenger, A., Fitzgerald, K., Godfrey, M., Gupta, D., McPherson, J., Malwadkar, P., Gupta, M., Bell, B., Doi, A.,

Jung, N., Li, X., Lynes, M. S., Brookes, E., Cherry, A. B., Demirbas, D., Tsankov, A. M., Zon, L. I., Rubin, L. L., Feinberg, A. P., Meissner, A., Cowan, C. A., and Daley, G. Q. (2015) A comparison of non-integrating reprogramming methods, *Nat Biotechnol 33*, 58-63.

11. Tsankov, A. M., Akopian, V., Pop, R., Chetty, S., Gifford, C. A., Daheron, L., Tsankova, N. M., and Meissner, A. (2015) A qPCR ScoreCard quantifies the differentiation potential of human pluripotent stem cells, *Nat Biotechnol 33*, 1182-1192.
12. Wesselschmidt, R. L. (2011) The teratoma assay: an in vivo assessment of pluripotency, *Human Pluripotent Stem Cells: Methods and Protocols* , 231-241.
13. Yaffe, M. P., Noggle, S. A., and Solomon, S. L. (2016) Raising the standards of stem cell line quality, *Nat Cell Biol 18*, 236-237.