Figure 6. Comparisons of prediction models. For the detailed definition of "M-pattern", see Material and Methods. Conceptual illustration of M-pattern appears in Figure 4. Potential I, osteogenic differentiation rate; potential II, adipogenic differentiation rate; potential III, chondrogenic differentiation rate; potential IV, PDT. Each matrix involving line plots consists of three columns, separated by dotted lines, representing differences among lots (Lots A, B, and C). In each column, horizontal axis represents passage numbers, from P2 on the left to P9 on the right. Upper number at the shoulder of each matrix indicates scaled error rate, i.e., the median value of prediction errors among all the samples, normalized by the experimental values. The lower number at the shoulder of each matrix indicates the correlation coefficient. Blue line plot represents the value of experimentally determined values. Red line plot represents the prediction values from the prediction models. Greater overlap between blue and red line plots and minimum differences across passages and lot differences corresponds to higher predictive performance, represented by lower scaled error rate and higher correlation coefficient.

exploration of the best LASSO model formula, it was possible to clearly detect the correlation of cytoskeleton-related and senescence-related genes with the decrease of potentials (Table S4–S7). Such correlation was supported by previous studies showing that the TGF-β signaling cascade links cellular quality collapse with morphological changes [20–22].

Compared to the successful predictions of adipogenic differentiation potential and PDT, the predictive performance for osteogenic and chondrogenic differentiation could not be increased by altering the modeling techniques. We believe that the main reason for this limitation on performance did not reflect a shortcoming of our method. Machine-learning performance relies heavily on the quality of training data. In this work, we are uncertain of the quality of our 'teacher signal' data, i.e., the converted data from conventional staining assays for evaluations of differentiation potential. In practice, the staining technique is usually used as only one aspect of differentiation confirmation, but is not commonly used for strict quantitative analysis. A growing body of evidence describes the quantitative use of immunohistochemical staining results in high-content analysis, analogous to the way in which we converted the staining results into numerical values. However, our machine-learning model results show that only Oil red [23], but not other staining values, results in excellent performances. In addition we have identified a few critical sources of experimental noise that might partially explain the comparatively poorer performance of osteogenic and chondrogenic differentiation. For example, in osteogenic staining with Alizarin red [24], small parts of stained cells tend to be ripped from the plate during the washing process, resulting in larger deviations within some wells. Similarly, in chondrogenic staining, the pellet size and its slice position greatly affect the staining level, resulting in larger deviations within some samples. Hence, the large difference between prediction models suggests that reproducibility and signal-to-noise ratio of staining results must be carefully examined in order for modeling to be effective. In other words, if one can introduce more stable staining, the prediction models should perform better.

Our expression data regarding other types of genes lead us to expect that effective gene combinations could be defined as new quality assessment markers (Fig. 4B, 4C). The interpretation of weights of LASSO regression models can provide insights regarding essential parameters that contributed to successful predictions (Table S4–S7). For prediction of potential I (osteogenic differentiation rate), a combination of morphological features from the whole pre-differentiation period (days 1–4), together with expression of the cytoskeleton-related genes (RAC1 [25] and RHOA [26]) and the early osteogenic marker SPP1 [osteopontin] [27], were weighted. The decision to weight these genes reflected previous reports of interactions between osteogenic marker genes and cytoskeleton genes [27]. Together, the selected morphological features support the commonly observed flat and expanded cellular morphology of hBMSCs, known as an indication of bone differentiation. For prediction of potential II (adipogenic differentiation rate), relative hole area (the morphological feature that describes the "roughness" of the cell surface) and inner radius (a reflection of polygonal and tentacle-like features in the cellular periphery) from the whole pre-differentiation period (days 1–4) were weighted. This result can be interpreted to mean that the continuous evolution of hBMSCs during the expansion period toward a jagged morphology is the signature of adipogenic potential. For prediction of potential III (chondrogenic differentiation rate), a shape factor (specifically, the roundness of cells) on day 1, but not throughout the whole period of expansion culture, was weighted. This can be interpreted to mean that very early

roundness of hBMSCs can indicate the potential for chondrogenesis. For prediction of differentiation potential before induction, morphogenic markers are sufficient because these data contain both time-course and multi-parametric information. However, in microarray experiments, we identified four genes related to osteogenic differentiation (ALPI, BMP2, BGLAP [osteocalcin], SSP1 [osteopontin]), three genes related to adipogenic differentiation (ADIPOQ [adiponectin], LEP [leptin], and LPL [lipoprotein lipase]), and one gene related to chondrogenic differentiation (ACAN [aggrecan]). From this background, these gene profiles were not sufficient to explain the difference in differentiation potential. As it happens, in prediction of potential I, PIK3CA and SPP1 [osteopontin] were useful markers even before differentiation induction; however, for the other two cell types, there were no critical marker genes. For prediction of potential IV (PDT of cells after repeated passages), both gene-expression profiles and morphological features were weighted. Cell cycle–related genes such as TP53 [p53] [28] and CDKN1A [p21] [29,30], actin-related genes such as ACTA2 [31] and IQGAP1 [32], and the cellular senescence–related gene CDKN2 [p16] [33–35] were heavily weighted in the best prediction models. Such gene selection reflects reports that indicate a correlation between senescence and cytoskeleton gene cascades in hBMSCs [19,36]. The weighted morphological features (total area, inner radius, and fiber length) reflect the common culture sense of regular size and slenderness during the expansion process are markers of active hBMSCs [19,37].

To confirm the utility of our proposed method in clinics, its adaptive performance in the context of the cellular diversity is an important criterion. Cellular diversity derived from patient diversity should be the first concern; we have previously investigated a new modeling scenario, designated as the "ongoing patient scenario", as one strategy for adapting to such diversity [17]. In this scenario, we proved that inclusion of new patients' own morphological features in prediction models can greatly enhance the prediction accuracy for new patients. Thus, this proposed method avoids attempting to adapt to all cellular variations arising due to patient diversity, which is ideal but impossible, and instead seeks to 're-train' the prediction model upon the arrival of each new patient, thereby allowing the system to adapt to new morphologies. Such re-training is feasible with our proposed regression models, and it is also feasible at various times in the clinic, because there are various opportunities to obtain images of primary cells before making critical potential predictions. Therefore, in this study, we attempted to optimize the adaptive performance of our model in the context of cellular diversity arising due to culture processes. In contrast to patient diversity, such culture process–derived diversity may expand during the cell production process, and may therefore require daily monitoring with non-invasive methods. Our results showed that even an extraordinarily diverse group of samples could be modeled feasibly using only morphological features. In addition, because our model performance is a summary of detailed cross-validation results combined with our ongoing patient scenario, we believe that our results provide a reliable performance benchmark reflecting robustness.

Throughout this work, we have successfully demonstrated models for predicting multiple future qualities of hBMSCs. Such models address an urgent need on the part of facilities that provide cell-based therapies. With similar objectives, Moghe's group and Gantenbein-Ritter's group have reported encouraging studies that reveal the effectiveness of multidimensional morphological parameter modeling in the evaluation of stem cell differentiation potentials [11–13]. However, in spite of the effectiveness of these

strategies, such quantitative and bioinformatic technologies are still not fully utilized in clinical cell therapies. To advance image-based cellular evaluation technologies as a reliable supportive option in regenerative medicine, it will be necessary to perform more studies that connect computational technology with stem cell biology. We believe our morphology-based modeling approach will contribute to new technological developments in regenerative medicine. It has been known that cell shape takes a part in the regulation of biological processes, such as proliferation and differentiation [27,38–40]. Although our main focus was to utilize such morphological information as "a signature of biological reflection" instead of investigating its meaning, our gene expression analysis and LASSO model interpretation have revealed the involvement of previously known cell shape-regulatory proteins, such as RHO-related proteins [27]. Therefore, we believe that our morphology-based modeling approach will contribute not only to the new technological developments in regenerative medicine, but also for deeper understanding of morphology effect in stem cell biology.

Materials and Methods

Cells and cell culture

Three lots of hBMSCs were purchased from Lonza (Walkersville, MD, USA): Lot A (lot number 8F3211, Black, Male, 18 year-old), Lot B (lot number 8F3434, Caucasian, Male, 22-year-old), and Lot C (lot number 8F3560, Hispanic, Female, 24-year-old). All cultures were maintained in MSCGM (Lonza) supplemented with BulletKit (Lonza), and sub-culture was performed with 0.025% trypsin solution (Life Technologies, Inc., Carlsbad, CA, USA). To expand the cell populations, each lot was continuously sub-cultured at 7–21-day intervals until the termination of cell growth. Passage timing was controlled by confluency. Specifically, passage was performed when confluency exceeded 80%, and we obeyed the supplier-recommended cell seeding density for Poietics[TM] human mesenchymal stem cells (Lonza). From all cell variants, P2–P9 samples were used. At each passage, cells were divided into three groups: passage sample (SEED group), pre-differentiation sample (PRE group), and differentiation sample (DIFF group). Sample I was subjected to an image-acquisition step. Sample R was subjected to RNA extraction for gene-expression profiling. The PRE groups represent the intact differentiation potentials of each hBMSC sample before the subsequent long-term differentiation process. The DIFF group was subdivided into three differentiation samples: Sample O for osteogenic differentiation, sample A for adipogenic differentiation, and sample C for chondrogenic differentiation. All differentiation cultures followed the protocol for Poietics Human Mesenchymal Stem Cells (Lonza). Briefly, samples were differentiated for 2 weeks (sample O), 3 weeks (sample A), or 4 weeks (sample C). See also Fig. 1 for the experimental scheme.

Image acquisition

For sample I, 10× magnification images at the exactly same positions (five fields at the center and four neighboring images separated by vertical displacements of 2.2 mm) were obtained every 24 hours for 4 days by phase-contrast microscopy on an X71 instrument regulated by an electric x-y stage (Olympus, Tokyo, Japan) (Fig. 1). The image data were gray-scale, 8-bit, 1360×1024 TIFF format. Focusing was semi-automatically defined by the original regulation journal run by MetaMorph (Molecular Devices, Sunnyvale, CA, USA). In total, 1,920 images (= 5 view fields×4 replicates of wells×3 lots×4 time points×8 passages) were stored in the image database.

Image processing

All phase-contrast microscopic images were processed using MetaMorph (Molecular device) with an original combination of image-processing filter sets (Figure S2 and Table S3) to measure morphological features of cells. In the binarization process, a single universal threshold was applied to all images. The universal threshold was defined as the threshold that provided the minimum error between manually determined cell number and the recognized total object counts (obtained using that threshold) among 30 images picked randomly from all lots and time points. After binarization, all individual objects in each image, consisting of cells and noise (non-cell objects), were measured using the integrated morphometry analysis function to measure nine morphological features: (1) breadth, (2) elliptical form factor, (3) fiber breadth, (4) fiber length, (5) hole area, (6) inner radius, (7) relative hole area, (8) shape factor, and (9) total area (detailed in Figure S3). In addition to morphological features, cellular object counts were added as the tenth feature. The morphological features were carefully selected in the MetaMorph measurement settings by clustering analysis, in order to eliminate problems arising from multi-colinearity. From this data, which consisted of object IDs and the corresponding morphological features, the noise (non-cell objects) was then removed by a noise-reduction algorithm developed in-house before subsequent analysis. The object morphological data obtained from each PRE sample summarized nearly 20 images consisting of 2,000–5,000 cellular objects. The morphological feature information representing these groups of objects was characterized by various statistical values, including average (AVE), quantile points (10, 25, 50, 75, and 90%), median (MED), interqurtile range (INT), robust skewness (SKEW), and robust kurtosis (KURT) (see detail in Fig. 5). Such statistics related to morphological values were obtained from four time points (24, 48, 72, and 96 hours). Six patterns of morphological feature sets (M-patterns 1–6) were designed to examine their effects on predictive performance (Fig. 5). M-pattern 1 (non-continuous average information), 40 features: AVE of nine morphology features at all four time points. M-pattern 2 (non-continuous quantile information), 184 features: five quantile points of nine morphology features at all four time points. M-pattern 3 (short continuous quantile information), 138 features: ratio of five quantile points of nine morphological features within three intervals between four time points. M-pattern 4 (non-continuous distribution pattern), 148 features: MED, INT, SKEW and KURT of nine morphological features at all four time points. M-pattern 5 (long continuous quantile information), 46 features: ratio of five quantile points of nine morphological features within one interval between 0 h and 96 h. M-pattern 6 (short non-continuous quantile information), 46 features: five quantile points of nine morphological features at 24 h only.

Quantitation of osteogenic differentiation rate

Calcium accumulation of differentiated samples was determined by Alizarin red staining [24] with some modifications. Cells were fixed with 70% ethanol for 1 hour, washed, and stained for 10 min with 40 mM alizarin red S solution (pH: 4.2). The quantitation of staining results was summarized by image-based measurements of the red pixels from 18 images (3 view fields×6 replicate wells, 1600×1200, color, RGB, jpg file) collected using a DP21 CCD camera (Olympus). The color image measurements were processed using MetaMorph (Molecular Devices) to extract the red pixels. The count of red pixels was taken as the quantitated value of the osteogenic differentiation rate.

— 220 —

Quantitation of adipogenic differentiation rate

Samples subjected to adipogenic differentiation were stained with Oil red S [23]. Briefly, cells were fixed with 4% paraformaldehyde (PFA) for 1 hour, washed, and stained for 30 min with Oil red S solution. The quantitation of staining results was summarized by image-based measurements of the area of stained droplets from 18 images (3 view fields×6 replicate wells, 1600×1200, color, RGB, jpg file) using a DP21 CCD camera. The stained lipid droplets were identified using MetaMorph (Molecular Devices), and their total pixel area was taken as the quantitated value of the adipogenic differentiation rate.

Quantitation of chondrogenic differentiation rate

Pellet samples differentiated into chondrocytes cultured under the chondrogenic differentiation condition were stained with Alcian blue [41]. Briefly, cells were fixed with 4% PFA for 1 hour, embedded in paraffin, and sliced into 20-μm sections. All sliced tissue samples were stained on the same day according to a conventional Alcian blue protocol. The quantitation of chondrogenic differentiation rate was obtained from image-based measurements of 1 image (1600×1200, color, jpg file) collected using a DP21 CCD camera. The area of Alcian blue staining was identified using MetaMorph (Molecular Device), and the total number of pixels was measured. To control for pellet size, the number of pixels staining positive for Alcian blue were normalized against the total pellet size (in pixels), and the resultant value was taken as the quantitated value of the chondrogenic differentiation rate.

Quantitation of population doubling time (PDT)

At each passage, the total cell number was counted to obtain the PDT [42].

Gene expression measurement

Total RNA was extracted using the RNAprotect Cell Reagent (Qiagen, Hilden, Germany) from sample R at confluence. RNA samples from P2–P9 were applied to the custom designed gene chip GenoPerl (Mitsubishi Rayon, Kyoto, Japan) (gene list in Table S2). The gene chip assay and analysis were carried out according to the manufacturer's protocols, and the data was used for prediction modeling. For global gene-expression analysis, expression levels of all genes were scaled by standard normalization between arrays and genes. The clustering heat map was created by Cluster 3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm#ctv) and Java Tree View (http://jtreeview.sourceforge.net) with some modifications.

Prediction model construction

The LASSO regression model was selected for modeling the relationships between morphological features and experimentally determined differentiation potentials. The detailed modeling process was previously described [17]. Briefly, the LASSO regression model is a penalized regression model that is widely used in the statistics and machine-learning literatures [43]. By using LASSO, one can find the linear combination of input features that best predict the teaching signal. LASSO tends to induce a sparse linear model, i.e., it can also select a set of input features that are useful for predictive purposes. The entire LASSO model-building process is automatic, including the model-selection process using leave-one-out cross-validation, and the relationships between morphological features and experimentally determined values can be represented as a simple and robust linear model. For model training, a total of 24 samples (P2–P9 in all three lots) were

used as the dataset. For input features, six patterns of morphological features (M-patterns 1–6) were assigned as non-invasively obtained information for cell-quality prediction. For teaching signal information, four types of experimentally determined differentiation potentials were used to train models: Potential I, osteogenic differentiation potential (Osteo), osteogenic differentiation rate determined by Alizarin red; potential II, adipogenic differentiation potential (Adipo), adipogenic differentiation rate determined by Oil red staining; potential III, chondrogenic differentiation potential (Chondro), chondrogenic differentiation rate determined by Alcian blue staining; and potential IV, population doubling time (PDT)[42] from the growth rate over continuous passages. For the relative comparison of prediction performance, a NULL prediction model (NULL model) was constructed. The NULL model is defined as the prediction model that sets the average value of all experimentally determined values as the threshold. This imitates the case of a poor prediction, and is easily established in practical experiments. Simply put, the NULL model is a negative-control model, similar to a random-guessing model restricted to one threshold value. Furthermore, gene-expression data were also assigned in the prediction model, as rivals of morphology-based cell-quality prediction models. From the custom microarray measurements, 69 gene-expression profiles were used as additional input features in the prediction modeling. Two types of prediction models were constructed: expression data for 69 genes in combination with the M-pattern 1 feature set, and the other was the expression data for 69 genes without morphological features. In total, 36 prediction models (9 types of input feature sets×4 types of differentiation potential predictions [3 lineages and population doubling time]) were constructed (Fig. 5).

Supporting Information

Figure S1 Conceptual illustration of usage and technological achievements of label-free morphology-based prediction of multiple differentiation potentials. A user can obtain three advantageous profits from our investigated method; (1) An early prediction, even from the images from the undifferentiation period to predict the final result after differentiation. Such prediction timing is designed to be fastened in this work, by the examination of the effect of early and sparse cellular images for future prediction. (2) An effective morphological feature conversion method, which can maximize the objective prediction of certain potential. Such morphological feature conversion method is comprehensively examined in this work, to reflect the meaning of heterogeneous nature of cells and their time-course changes by various ideas of morphological feature calculations. (3) A multiple simultaneous prediction for same image. In our method, four types of potential prediction model are constructed, and provide results at the same time for one image. Such paralleled prediction concept enables "overlapping" multiple evaluations of cells with non-invasive manner.
(TIF)

Figure S2 Schematic procedure of image processing and data processing. (A) The procedure listed as Filter 1–4 in Table S3 is illustrated. (B) The procedure listed as Filer 5–6 in Table S3 is illustrated. Especially, the illustration describes the detail of cell measurement and their data processing scheme followed by Filter 6 processing. As shown in the figure, all cells in the images are measured as data consist of "group of cells", and their distribution is used to calculate statistic values to describe such "group of cells". Through the process, ells are measured individually by morphological indices; however our final morpho-

logical features reflect the information of "group of cells". In other words, our morphological features contain the information of "heterogeneity" of "group of cells", which strengthen our prediction model.
(TIF)

Figure S3 Detailed definition of the basic nine morphological features used for cell measurements. Formulas and schematic illustration of morphological features are presented in detail. The nine features are carefully selected to represent independent information, with the aspect of low-correlating parameters, for stabilizing prediction models which utilize morphological features from label-free phase contrast images.
(TIF)

Figure S4 Low-magnification images of cells in Figure 2A. The low-magnification images provides overall image of cellular morphological profile and its distribution.
(TIF)

Table S1 Dataset profile used for prediction model construction.
(XLSX)

Table S2 Gene list in custom designed gene chip microarray.
(XLSX)

Table S3 Image processing scheme.
(XLSX)

Table S4 Parameters and weights explored to contribute in prediction models for Potential I.
(XLSX)

Table S5 Parameters and weights explored to contribute in prediction models for Potential II.
(XLSX)

Table S6 Parameters and weights explored to contribute in prediction models for Potential III.
(XLSX)

Table S7 Parameters and weights explored to contribute in prediction models for Potential IV.
(XLSX)

Acknowledgments

Author Contributions

References

1. Arthur A, Zannettino A, Gronthos S (2009) The therapeutic applications of multipotential mesenchymal/stromal stem cells in skeletal tissue repair. J Cell Physiol 218: 237–245.
2. Braude P, Minger S, Warwick R (2005) Stem cell therapy: hope or hype? Safety and quality must be assured before this treatment can really benefit patients. Br Med J 330: 1159–1160.
3. Wang Y, Han ZB, Song YP, Han ZC (2012) Safety of mesenchymal stem cells for clinical application. Stem Cells Int 2012: 652034.
4. Werbowetski-Ogilvie T, Bosse M, Stewart M, Schnerch A, Ramos-Mejia V, et al. (2009) Characterization of human embryonic stem cells with features of neoplastic progression. Nat Biotechnol 27: 91–97.
5. Settleman J (2004) Tension precedes commitment - even for a stem cell. Mol Cell 14: 148–150.
6. Docheva D, Padula D, Popov C, Mutschler W, Clausen-Schaumann H, et al. (2008) Researching into the cellular shape, volume and elasticity of mesenchymal stem cells, osteoblasts and osteosarcoma cells by atomic force microscopy. J Cell Mol Med 12: 537–552.
7. Zhang L, Yu J, Pan H, Hu P, Hao Y, et al. (2007) Small molecule regulators of autophagy identified by an image-based high-throughput screen. Proc Natl Acad Sci USA 104: 19023–19028.
8. Carpenter A, Jones T, Lamprecht M, Clarke C, Kang I, et al. (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7: R100.
9. Futamura Y, Kawatani M, Kazami S, Tanaka K, Muroi M, et al. (2012) Morphobase, an encyclopedic cell morphology database, and its use for drug target identification. Chem Biol 19: 1620–1630.
10. Misselwitz B, Strittmatter G, Periaswamy B, Schlumberger MC, Rout S, et al. (2010) Enhanced CellClassifier: a multi-class classification tool for microscopy images. BMC Bioinformatics 11.
11. Treiser MD, Yang EH, Gordonov S, Cohen DM, Androulakis IP, et al. (2010) Cytoskeleton-based forecasting of stem cell lineage fates. Proc Natl Acad Sci U S A 107: 610–615.
12. Kim JJ, Vega SL, Moghe PV (2013) A high content imaging-based approach for classifying cellular phenotypes. Methods in molecular biology 1052: 41–48.
13. Vega SL, Liu E, Patel PJ, Kulesa AB, Carlson AL, et al. (2012) High-content imaging-based screening of microenvironment-induced changes to stem cells. J Biomol Screen 17: 1151–1162.
14. Seiler C, Gazdhar A, Reyes M, Benneker LM, Geiser T, et al. (2012) Time-lapse microscopy and classification of 2D human mesenchymal stem cells based on cell shape picks up myogenic from osteogenic and adipogenic differentiation. J Tissue Eng Regen Med.
15. Friedenstein A, Gorskaja U, Kulagina N (1976) Fibroblast precursors in normal and irradiated mouse hematopoietic organs. Exp Hematol 4: 267–274.
16. Caplan AI, Bruder SP (2001) Mesenchymal stem cells: building blocks for molecular medicine in the 21st century. Trends in molecular medicine 7: 259–264.
17. Matsuoka F, Takeuchi I, Agata H, Kagami H, Shiono H, et al. (2013) Morphology-based prediction of osteogenic differentiation potential of human mesenchymal stem cells. PloS ONE 8: e55082.
18. Konstantinov B, Zhou W, Golini F, Hu W-S (1994) Expert-system in the control of animal-cell culture processes - potentials, functions, and perspectives. Cytotechnology 14: 233–246.
19. Røsland GV, Svendsen A, Torsvik A, Sobala E, McCormack E, et al. (2009) Long-term cultures of bone marrow derived human mesenchymal stem cells frequeantly undergo spontaneous malignant transformation. Cancer Res 69: 5331–5339.
20. Ito T, Sawada R, Fujiwara Y, Tsuchiya T (2008) FGF-2 increases osteogenic and chondrogenic differentiation potentials of human mesenchymal stem cells by inactivation of TGF-beta signaling. Cytotechnology 56: 1–7.
21. Wang DJ, Park JS, Chu JSF, Krakowski A, Luo KX, et al. (2004) Proteomic profiling of bone marrow mesenchymal stem cells upon transforming growth factor beta 1 stimulation. J Biol Chem 279: 43725–43734.
22. Massague J, Xi Q (2012) TGF-beta control of stem cell differentiation genes. FEBS letters 586: 1953–1958.
23. Kinkel A, Fernyhough M, Helterline D, Vierck J, Oberg K, et al. (2004) Oil red-O stains non-adipogenic cells: a precautionary note. Cytotechnology 46: 49–56.
24. Gregory C, Gunn W, Peister A, Prockop D (2004) An Alizarin red-based assay of mineralization by adherent cells in culture: comparison with cetylpyridinium chloride extraction. Anal Biochem 329: 77–84.
25. Jordan P, Brazao R, Boavida MG, Gespach C, Chastre E (1999) Cloning of a novel human Rac1b splice variant with increased expression in colorectal tumors. Oncogene 18: 6835–6839.
26. Jaffe AB, Hall A (2005) Rho GTPases: biochemistry and biology. Annu Rev Cell Dev Biol 21: 247–269.
27. McBeath R, Pirone D, Nelson C, Bhadriraju K, Chen C (2004) Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. Dev Cell 6: 483–495.
28. Zuckerman V, Wolyniec K, Sionov R, Haupt S, Haupt Y (2009) Tumour suppression by p53: the importance of apoptosis and cellular senescence. J Pathol 219: 3–15.
29. Motaln H, Schichor C, Lah TT (2010) Human mesenchymal stem cells and their use in cell-based therapies. Cancer 116: 2519–2530.
30. Ong CC, Jubb AM, Haverty PM, Zhou W, Tran V, et al. (2011) Targeting p21-activated kinase 1 (PAK1) to induce apoptosis of tumor cells. Proc Natl Acad Sci USA 108: 7177–7182.
31. Comer K, Dennis P, Armstrong L, Catino J, Kastan M, et al. (1998) Human smooth muscle alpha-actin gene is a transcriptional target of the p53 tumor suppressor protein. Oncogene 16: 1299–1308.

32. Weissbach L, Settleman J, Kalady M, Snijders A, Murthy A, et al. (1994) Identification of a human RASGAP-related protein containing calmodulin-binding motifs. J Biol Chem 269: 20517–20521.

33. Jacobs J, Kieboom K, Marino S, DePinho R, van Lohuizen M (1999) The oncogene and polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus. Nature 397: 164–168.

34. Janzen V, Forkert R, Fleming HE, Saito Y, Waring MT, et al. (2006) Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16(INK4a). Nature 443: 421–426.

35. Shibata KR, Aoyama T, Shima Y, Fukiage K, Otsuka S, et al. (2007) Expression of the p16INK4A gene is associated closely with senescence of human mesenchymal stem cells and is potentially silenced by DNA methylation during in vitro expansion. Stem Cells 25: 2371–2382.

36. Bin Noh H, Ahn H, Lee W, Kwack K, Do Kwon Y (2010) The molecular signature of in vitro senescence in human mesenchymal stem cells. Genes Genomics 32: 87–93.

37. Ito T, Sawada R, Fujiwara Y, Seyama Y, Tsuchiya T (2007) FGF-2 suppresses cellular senescence of human mesenchymal stem cells by down-regulation of TGF-beta 2. Biochem Biophys Res Commun 359: 108–114.

38. Chen CS, Mrksich M, Huang S, Whitesides GM, Ingber DE (1997) Geometric control of cell life and death. Science 276: 1425–1428.

39. Roskelley CD, Desprez PY, Bissell MJ (1994) Extracellular Matrix-Dependent Tissue-Specific Gene-Expression in Mammary Epithelial-Cells Requires Both Physical and Biochemical Signal-Transduction. Proc Natl Acad Sci USA 91: 12378–12382.

40. Watt FM, Jordan PW, Oneill CH (1988) Cell-Shape Controls Terminal Differentiation of Human Epidermal-Keratinocytes. Proc Natl Acad Sci USA 85: 5576–5580.

41. Steedman H (1950) Alcian blue 8 GS - A new stain for mucin. Q J Microsc Sci 91: 477–479.

42. Ben Azouna N, Jenhani F, Regaya Z, Berraeis L, Ben Othman T, et al. (2012) Phenotypical and functional characteristics of mesenchymal stem cells from bone marrow: comparison of culture using different media supplemented with human platelet lysate or fetal bovine serum. Stem Cell Res Ther 3: 6–10.

43. Tibshirani R (1996) Regression shrinkage and selection via the Lasso. J Roy Stat Soc Series B Stat Methodol 58: 267–288.

— 223 —

Cyclin D2 Promotes the Proliferation of Human Mesenchymal Stem Cells

Ken Kono, Shingo Niimi and Rumi Sawada*

Division of Medical Devices, National Institute of Health Sciences, Japan

Abstract

Background: Human mesenchymal stem cells (hMSCs) hold promise for use in cell-based therapies and tissue engineering. Although hMSCs are thought to be stable *ex vivo*, it is possible that they undergo an undesirable transformation to a phenotype of unlimited proliferation during *ex vivo*. In this study, we searched for the factor required for unlimited proliferation of hMSCs.

Methods: Changes in gene expression were evaluated between hMSCs and Ewing's sarcoma cell lines, which may be derived from hMSCs, using GeneChip Human Genome U133 plus 2.0 Array. A gene up-regulated by at least 10-fold in Ewing's sarcoma cell lines, Cyclin D2, was overexpressed in hMSCs by a lentiviral vector.

Results: Overexpression of Cyclin D2 in hMSCs altered cell morphology and promoted cell proliferation. Expression of transforming growth factor-β2 (TGF-β2), which induces senescence in hMSCs, was down-regulated in Cyclin D2-overexpressing hMSCs. Furthermore, Gene Ontology analysis revealed that Cyclin D2 overexpression activated expression of genes associated with proliferation and interphase.

Conclusions: Cyclin D2 promotes hMSC proliferation and is a candidate biomaker for hMSC transformation.

Keywords: hMSCs; Ewing's sarcoma; Cyclin D2; Cell proliferation

Introduction

Mesenchymal stem cells (MSCs) self-replicate and differentiate into a variety of cell types such as osteoblasts, chondrocytes, adipocytes, and smooth muscle cells [1-5]. These capacities have made MSCs useful in studies of bone and cartilage regeneration [6-8]. One of the sources of human MSCs (hMSCs) is adult bone marrow, although they occur at a rate of one per one-hundred-thousand nucleated cells [6], and the available volume of bone marrow is limited. To secure the numbers of hMSCs required for tissue regeneration, the cell must be expanded *ex vivo*. Although hMSCs are stable *ex vivo*, it is possible that they undergo transformation to an unlimited proliferation phenotype during expansion.

Previous studies have demonstrated that Ewing's sarcoma is derived from MSCs [9-12]. Ewing's sarcoma is a malignancy that primarily affects children and young adults, with a peak incidence between the ages of 14 and 20 years. It arises mainly in bone and less commonly in soft tissues. The t(11;22)(q24;q12) chromosomal translocation generating EWS-FLI-1 fusion gene is found in 85% of cases [13]. EWS-FLI-1 knockdown inhibits cell proliferation in Ewing's sarcoma cells [14,15]. Thus, EWS-FLI-1 expression is believed to play a key role in Ewing's sarcoma development. However, EWS-FLI-1 expression does not transform normal murine and human fibroblasts [16,17], suggesting EWS-FLI-1 promotes malignant transformation in selective cells.

Several reports have demonstrated that EWS-FLI-1 expression transforms murine MSCs; indeed, tumors form when these cells are injected into immunodeficient mice [9,12]. In contrast, EWS-FLI-1 expression in hMSCs does not accelerate cell proliferation and transformation (10). EWS-FLI-1 expression in hMSCs induces a gene expression profile that closely mimics that of Ewing's sarcoma [9-11] without affecting proliferation. Therefore, MSCs are thought to be the origin of Ewing's sarcoma, but because EWS-FLI-1 alone cannot transform hMSCs, we believe other factors are required for transformation.

The most important safety concern when using hMSCs in cell-

based therapies and tissue engineering is the occurrence of unlimited proliferation during *ex vivo* culture. To identify the factors required for unlimited hMSC proliferation, we compared the gene expression profiles of hMSCs and Ewing's sarcoma cell lines and found that Cyclin D2 expression was extremely high in the Ewing's sarcoma cell lines. Overexpression of Cyclin D2 promotes proliferation of hMSCs, suggesting that Cyclin D2 is a candidate biomaker for hMSC transformation.

Materials and Methods

Cell culture

hMSCs derived from bone marrow were purchased from Lonza (Walkersville, MD) and cultured in MSCGM BulletKit, a mesenchymal stem cell basal medium with mesenchymal cell growth supplement, L-glutamine, and gentamycin/amphotericin-B (Lonza Walkersville, MD). Ewing's sarcoma cell lines (Hs 822.T, Hs 863.T, RD-ES, and SK-ES-1) were purchased from American Type Culture Collection (ATCC; Manassas, VA). Hs 822.T and Hs 863.T were cultured in Dulbecco's Modified Eagle's medium (DMEM; Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS; Intergene). RD-ES was cultured in RPMI-1640 medium (Gibco) supplemented with 10% FBS. SK-ES-1 was cultured in McCoy's 5a medium modified (Gibco) supplemented with 15% FBS. 293T (human kidney; ATCC) was cultured in DMEM supplemented with 10% FBS.

*Corresponding authors: Rumi Sawada, Division of Medical Devices, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan, Tel: 81-3-3700-1487; Fax: 81-3-3700-9196; E-mail: rsawada@nihs.go.jp

Microarray analysis

Total RNA was extracted from hMSCs and Ewing's sarcoma cell lines with the RNeasy Mini Kit (QIAGEN, Valencia, CA). Total RNA quantity and quality were assessed on an Agilent 2100 Bio-analyzer (Agilent, Santa Clara, CA); 100 ng total RNA was used to generate biotin-modified amplified RNA (aRNA) with GeneChip 3′IVT Express Kit (Affymetrix, Santa Clara, CA). Reverse transcription (RT) of first-strand complementary DNA (cDNA) with the T7 promoter sequence was performed with the T7 oligi(dT) primer. Second-strand cDNA synthesis was used to convert the single-stranded cDNA into a double-stranded DNA template by using DNA polymerase and RNase H to simultaneously degrade the RNA and synthesize second-strand cDNA. In vitro transcription of biotin-modified aRNA with IVT Labeling Master Mix generated multiple copies of biotin-modified aRNA from the double-stranded cDNA templates. The aRNA was purified and quantified; after fragmentation, it was hybridized to GeneChip Human Genome U133 Plus 2.0 Array (Affymetrix). The arrays were stained with phycoerythrin and washed at the GeneChip Fluidics station 450 (Affymetrix). The microarrays were scanned and data extracted using GeneChip scanner 3000 7G (Affymetrix); image analysis was performed using the Affymetrix GeneChip Command Console Software and digitized with the Affymetrix Expression Console.

Data processing and pathway analysis

Data analysis was performed with GeneSpring GX 11.0 software (Agilent Technologies, Santa Clara, CA). Raw data were normalized to the 50th percentile per chip and the median per gene. Differentially expressed genes were analyzed using Ingenuity Pathway Analysis (IPA) 9.0 (Ingenuity Systems, Redwood City, CA). Fisher's exact test was used to calculate a P-value. Activation z-score was calculated as a measure of functional and translational activation in Networks and Upstream regulators analysis. An absolute z-score >2 was considered significant.

Real-time RT-PCR

Total RNA was reverse-transcribed with SuperScript III First-Strand Synthesis System for RT-PCR (Life Technologies Co., Carlsbad, CA). Real-time RT-PCR was performed with LightCycler Fast Start DNA Master SYBR Green I (Roche Applied Science, Basel, Switzerland) in a Roche LightCycler instrument (software version 4.0). mRNA expression was normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The primers for Cyclin D2 and p16 were 5′-TACTTCAAGTGCGTGCAGAAGGAC-3′ and 5′-TCCCA-CACTTCCAGTTGCGATCAT-3′ (Cyclin D2) and 5′-CACTCAC-GCCCTAAGC-3′ and 5′-GCAGTGTGACTCAAGAGAA-3′ (p16). The primers for transforming growth factor-β2 (TGF-β2) and GAPDH were from Light Cycler Primer Sets (Search LC GmbH, Heidelberg, Germany).

Cloning and expression of Cyclin D2

Cyclin D2 cDNA was amplified by RT-PCR of mRNA extracted from SK-ES-1 using 5′-GAATTCGCCACCATGGAGCTGCTGT-GCCACGAGG-3′ (forward; EcoR I site underlined) and 5′-CTC-GAGTCACAGGTCGATATCCCGCACG-3′ (reverse; Xho I site underlined). The amplified products were cloned into pTA2 (ToYoBo, Osaka, Japan) and verified by sequencing. The verified Cyclin D2 cDNA was cloned into the EcoR I and Xho I sites of pLVSIN-CMV Pur (TaKaRa, Shiga, Japan). Lentiviral vector was prepared with the Lenti-XTM Packaging System (TaKaRa) according to manufacturer protocols.

Viral infection

hMSCs were infected with the lentiviral vector containing Cyclin D2 (hMSCs/CyclinD2) or empty vector (hMSCs/Empty) at 37°C for 24 h. Infected cells were selected with 1 μg/mL puromycin for 14 days and the bulk of the resistant cells was used in subsequent experiments.

Western blotting

hMSCs/CyclinD2 and hMSCs/Empty were lysed in RIPA buffer (Wako, Osaka, Japan). Cyclin D2 was separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a membrane (Immunobilon-P^{SQ}; Millipore, Billerica, MA). Blots were blocked and probed overnight at 4°C with a mouse monoclonal antibody against Cyclin D2 (MBL, Nagoya, Japan). Blots were incubated with peroxidase-conjugated anti-mouse IgG (Abcam) and bound antibodies were visualized with Chemi-Lumi One Super Chemiluminescence (Nacalai Tesque, Kyoto, Japan).

Cell proliferation

Proliferation of hMSCs/CyclinD2 and hMSCs/Empty was measured with the TetraColor ONE reagent (Seikagaku Co., Tokyo, Japan). Cultures were incubated for 2 h in medium containing the reagent. Absorbance was read at 450 nm (reference at 600 nm) on a plate reader (SH-9000, Corona Electric Co., Ibaraki, Japan).

Results

Cyclin D2 expression in Ewing's sarcoma cell lines versus hMSCs

To identify the factors required for proliferation of hMSCs, we compared the gene expression profiles of hMSCs and four Ewing's sarcoma cell lines (Hs 822.T, Hs 863.T, RD-ES, and SK-ES-1) (Figure 1a). Hs 822.T and Hs 863.T had similar expression profiles, as did RD-ES and SK-ES-1. The expression profiles of Hs 822.T and Hs 863.T were more similar to those of hMSCs than to those of the other Ewing's sarcoma (RD-ES and SK-ES-1). Therefore, we first compared the expression profiles of hMSCs, Hs 822.T, and Hs 863.T. We identified 44 genes that differed by at least 10-fold between hMSCs and Ewing's sarcoma cell lines (data not shown). These were narrowed to 9 genes by selecting genes that also differed from hMSCs by more than 10-fold in RD-ES and SK-ES-1 (Table 1). CCND2 (Cyclin D2) stood out in this group of 9 genes, because it represents a family of key cell-cycle regulators. Indeed, aberrant expression of
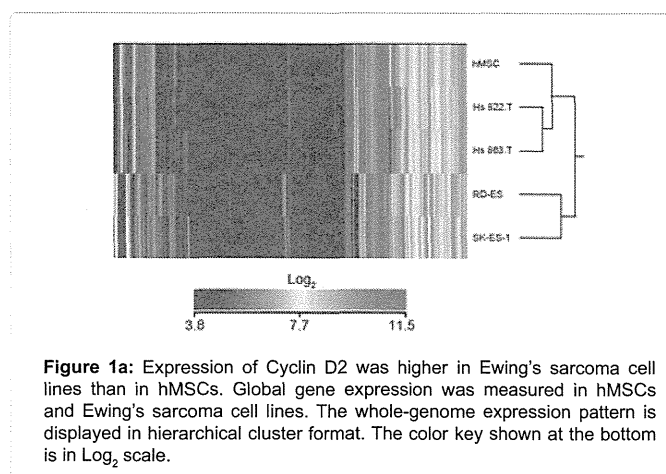

Figure 1a: Expression of Cyclin D2 was higher in Ewing's sarcoma cell lines than in hMSCs. Global gene expression was measured in hMSCs and Ewing's sarcoma cell lines. The whole-genome expression pattern is displayed in hierarchical cluster format. The color key shown at the bottom is in Log$_2$ scale.
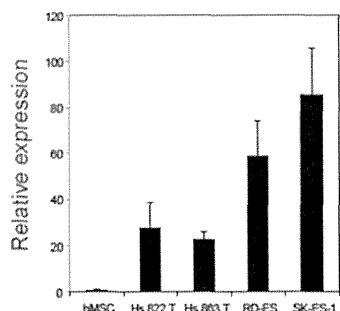
Figure 1b: Expression of Cyclin D2 was higher in Ewing's sarcoma cell lines than in hMSCs. Cyclin D2 expression in hMSCs and Ewing's sarcoma cell lines was determined by real-time RT-PCR with values normalized to GAPDH. Results are plotted relative to hMSCs. Data from triplicate samples (means ± SD) are shown.
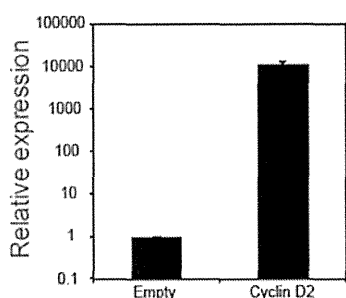

Figure 2a: Over expression of Cyclin D2 promoted hMSC proliferation. At 14 days after infection when puromycin selection was completed, Cyclin D2 mRNA in hMSCs/CyclinD2 and hMSCs/Empty was determined by real-time RT-PCR and normalized to GAPDH. Results are plotted relative to hMSCs/ Empty.
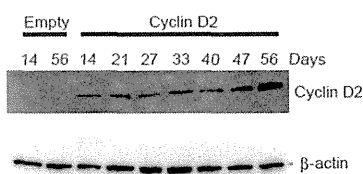

Figure 2b: Overexpression of Cyclin D2 promoted hMSC proliferation. Cyclin D2 protein expression was visualized by western blotting at the indicated days after infection. Equal amounts of protein were loaded in each lane.

Cyclin D2 has been associated with tumor progression in many tumor types [18-21]. We measured Cyclin D2 mRNA expression in hMSCs and Ewing's sarcoma cell lines by real-time PCR and confirmed its extreme induction in Ewing's sarcoma (Figure 1b). Hs 822.T and Hs 863.T exhibited similar Cyclin D2 expression levels; expression was higher in RD-ES and SK-ES-1.

Overexpression of Cyclin D2 promoted hMSC proliferation

We transduced the Cyclin D2 gene into hMSCs by using a lentiviral vector; an empty vector served as a negative control. At 14 days after infection when puromycin selection was completed, Cyclin D2 mRNA expression in hMSCs infected with lentiviral vector containing Cyclin D2 (hMSCs/CyclinD2) was about 10,000-fold higher than in hMSCs infected with empty vector (hMSCs/Empty) (Figure 2a). Cyclin D2 was stably expressed in hMSCs at least 56 days after infection (Figure 2b). Next, we observed the cell morphology of the hMSCs/CyclinD2. Phase contrast microscopy revealed the normal fibroblast-like morphology of hMSCs/Empty (Figure 2c) and smaller spread areas in hMSCs/ CyclinD2 (Figures 2c and 2d). Furthermore, proliferation of hMSCs/ CyclinD2 was faster than that of hMSCs/Empty, indicating that overexpression of Cyclin D2 promoted hMSC proliferation. However, the proliferation slowed over time and did not result in unlimited proliferation (data not shown).

TGF-β2 expression was down-regulated in hMSCs/CyclinD2, but p16 expression was not

To investigate the effects of overexpression of Cyclin D2 on the cell cycle, we examined the change in cell cycle-associated gene expression over time (p16, p21, Bmi1, TGF-β1, and TGF-β2). We did not detect significant differences in expression of p21, Bmi1, and TGF-β1 between hMSCs/CyclinD2 and hMSCs/Empty (data not shown). However, TGF-β2 expression was lower in hMSCs/CyclinD2 than in hMSCs/ Empty (Figure 3a). In addition, the increase in TGF-β2 expression during culture was suppressed in hMSCs/CyclinD2 compared with hMSCs/Empty. In contrast, the increasing rate of p16 expression in hMSCs/CyclinD2 was higher than in hMSCs/Empty, although expression in both cell types was comparable 14 days after infection (Figure 3b).

Overexpression of Cyclin D2 altered the expression of genes associated with cell proliferation and interphase

Total RNA was extracted from hMSCs/CyclinD2 and hMSCs/ Empty 14 days after infection and analyzed by DNA microarray, which identified 690 genes that were differentially expressed by at least 2-fold between hMSCs/CyclinD2 and hMSCs/Empty

Gene Symbol	Entrez Gene Name	Fold Change			
		Hs822.1	Hs863 T	RD-ES	SK-ES-1
BIEND5	BEN domain containing 5	-16.720	-14.819	27.302	26.690
CCND2	Cyclin D2	38.078	37.702	67.745	247.690
FBLN1	Fibulin1	-18.992	-24.363	-36.755	-198.523
HAPLN 1	Hyaluronan and and proteo glycan link protein 1	-22.460	-11.738	-26.674	-110.084
1GF2BP1	Insulin-Re growth factor 2 mRNA binding protein 1	22.847	22.358	288.588	207.475
SLC24A3	Solute carrier may 24 (sodium/potassium calcum exchanger), member 3	38.867	75.637	324.331	173.842
SSTR1	Somatostain receptor 1	79.594	14.840	22.406	28.783
STEAP4	STEAP family member 4	-24.490	-18.837	-21.747	-33.153
TMEFF2	Transmembrane protein with EGF-like and two follistatin-ike domain 2	37.577	10.463	34.418	197.834

Table 1: Expression of Cyclin D2 was higher in Ewing's sarcoma cell lines than in hMSCs. The nine genes differentially expressed by at least 10-fold between hMSCs and the four Ewing's sarcoma cell lines.
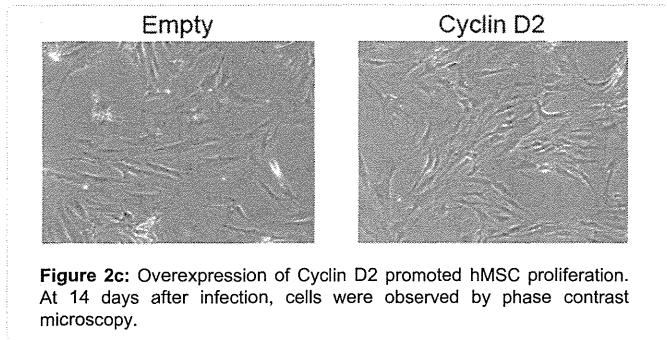
Figure 2c: Overexpression of Cyclin D2 promoted hMSC proliferation. At 14 days after infection, cells were observed by phase contrast microscopy.
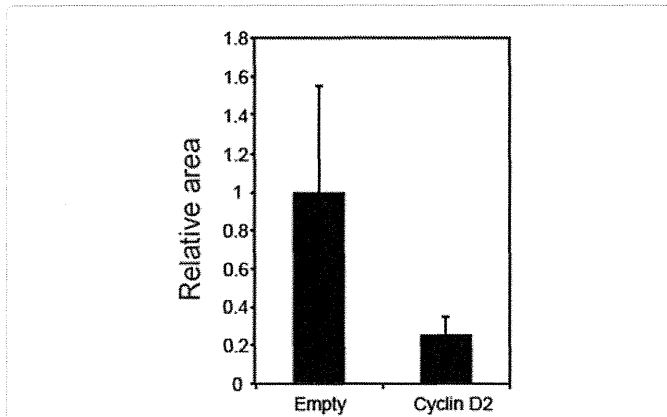

Figure 2d: Overexpression of Cyclin D2 promoted hMSC proliferation. Fifteen random hMSCs/Empty and hMSCs/CyclinD2 cells were measured with Image J 1.43t software (NIH, USA). Values were plotted relative to the averaged area of hMSCs/Empty.
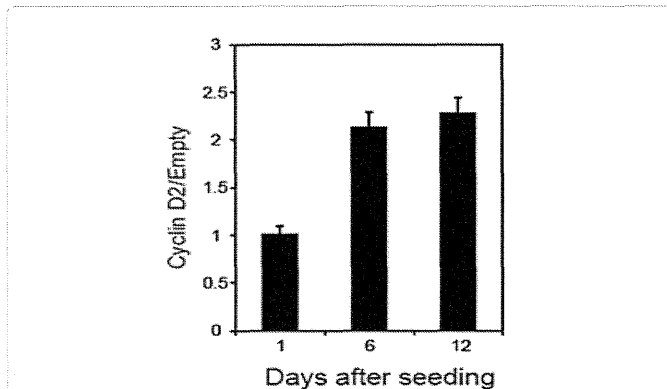

Figure 2e: Overexpression of Cyclin D2 promoted hMSC proliferation. At 14 days after infection, equal numbers of cells were seeded in 24-well plates and proliferation was determined with Tetra Color One 1, 6, and 12 days after seeding. The optical density (Cyclin D2/Empty) of triplicate samples is shown.

(Figure 4a). Gene ontology (GO) analysis revealed these genes are associated with movement, development, growth and proliferation, cell cycle, and intercellular signaling and interactions (Table 2). Specific predictions indicated that proliferation and interphase are activated in hMSCs/CyclinD2. The induced genes that are associated with proliferation and interphase are listed in tables 3 and 4; in summary, 94 of 186 genes and 19 of 50 genes exhibited expression shifts consistent with increased in proliferation and interphase, respectively.

Discussion

Although EWS-FLI-1 expression transformed murine MSCs, expression in hMSCs did not promote cell proliferation. In this study, we found that Cyclin D2 expression was extremely high in the Ewing's sarcoma cell lines and overexpression of Cyclin D2 in hMSCs promoted cell proliferation. GO analysis also predicted that cell proliferation and interphase were activated by overexpression of Cyclin D2.

Cyclin D2 is a member of the family of D-type cyclins that mediate cell cycle regulation, differentiation, and oncogenic transformation [22,23]. D-type cyclins inactivate retinoblastoma (Rb) by phosphorylation, inducing release of E2F. Free E2F activates genes involved in the activation and maintenance of DNA synthesis. Thus, overexpression of Cyclin D2 generally has growth-promoting effects. Consistent with this notion, overexpression of Cyclin D2 in HeLa
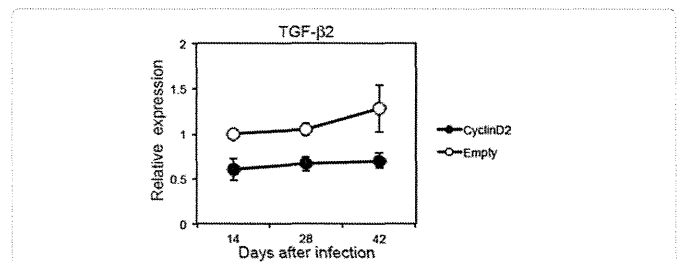

Figure 3a. TGF-β2 expression over time. TGF-β2 mRNA expression was determined by real-time RT-PCR at 14, 28, and 42 days after infection. Results are plotted relative to hMSCs/Empty at 14 days after infection. A representative of three independent experiments is shown

Figure 3b. p16 expression over time. p16 mRNA expression was determined by real-time RT-PCR at 14, 28, and 42 days after infection. Results are plotted relative to hMSCs/Empty at 14 days after infection. A representative of three independent experiments is shown.
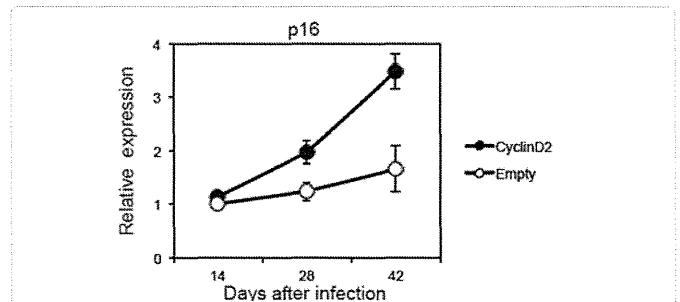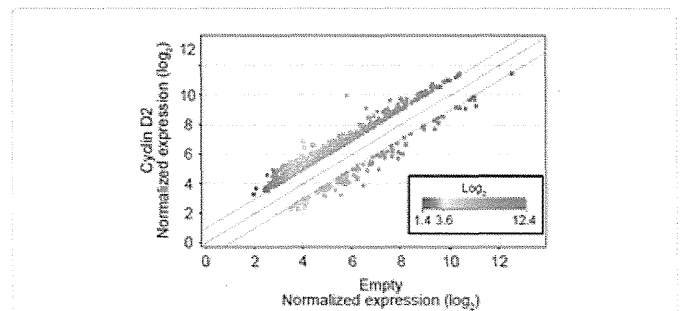

Figure 4a. Global gene expression in hMSCs/CyclinD2. Global gene expression patterns were compared in hMSCs/CyclinD2 and hMSCs/Empty. Genes differentially expressed by at least 2-fold are shown (690 genes). The color key shown at the bottom is in Log_2 scale

Top 5 functional category	Function annotation	p-Value	Activation z-score	No. of genes
Cellular Movement	cell movement of prostate cancer	7.94E-06	0.918	14
	migration of prostate cancer cells	1.54E-05	1.339	12
	recruitment of cells	6.57E-05	1.55	28
Cellular Development	proliferation of tumor cell lines	9.04E-06	1.15	92
	proliferation of cancer cells	1.96E-05	1.302	30
	differentiation of connective tissues	4.35E-05	-0.379	41
	proliferation of tumor cells	4.61E-05	1.036	37
Cellular Growth and Proliferation	proliferation of tumor cell lines	9.04E-06	1.15	92
	proliferation of cancer cells	1.96E-05	1.302	30
	proliferation of tumor cells	4.61E-05	1.036	37
	proliferation of cells	9.27E-05	3.142	186
Cell Cycle	interphase	1.34E-05	2.19	50
Cell-To-Cell Signaling and Interaction	recruitment leukocytes	7.13E-05	1.159	26

Table 2: Global gene expression in hMSCs/CyclinD2. Gene ontology (GO) analysis of the 690 genes was performed with Ingenuity Pathway Analysis (IPA) 9.0. Top five functional categories and the specified categories are listed. An absolute z-score >2 was considered as significant.

ID	Genes in dataset	Fold Change	Prediction (based on expression direction)	ID	Genes in dataset	Fold Change	Prediction (based on expression direction)	ID	Genes in dataset	Fold Change	Prediction (based on Expression direction)
209292_at	ID4	3.609	Increased	201195_s_at	SLC7A5	2.083	Increased	203083_at	THBS2	2.096	Decreased
206271_at	TLR3	2.023	Increased	222749_at	SUFU	-3.251	Increased	203468_at	CDK10	2.228	Decreased
203373_at	SOCS2	3.513	Increased	204508_s_at	CA12	3.237	Increased	1552721_a_at	FGF1	-2.054	Decreased
206649_s_at	TFE3	2.069	Increased	203764_at	DLGAP5	2.188	Increased	204052_s_at	SFRP4	2.249	Decreased
201292_at	TOP2A	2.434	Increased	214012_at	ERAP1	5.288	Increased	205548_s_at	BTG3	2.073	Decreased
204766_s_at	NUDT1	2.158	Increased	204614_at	SERPINB2	-2.048	Increased	242989_at	STRN	2.145	Decreased
209321_s_at	ADCY3	2.329	Increased	201416_at	SOX4	3.11	Increased	213905_x_at	BGN	2.09	Decreased
206693_at	IL7	2.399	Increased	214581_x_at	TNFRSF21	-2.049	Increased	228780_at	POU3F3	-2.721	Decreased
205345_at	BARD1	2.029	Increased	223570_at	MCM10	2.545	Increased	205016_at	TGFA	-3.032	Decreased
226377_at	NFIC	2.026	Increased	219743_at	HEY2	2.13	Increased	1569791_at	STK4	2.396	Decreased
212148_at	PBX1	2.514	Increased	209970_x_at	CASP1	3.111	Increased	216008_s_at	ARIH2	2.331	Decreased
202213_s_at	CUL4B	2.093	Increased	202684_s_at	RNMT	2.503	Increased	221577_x_at	GDF15	2.143	Decreased
224954_at	SHMT1	2.105	Increased	242979_at	IRS1	2.329	Increased	202153_s_at	NUP62	2.009	Decreased
209960_at	HGF	2.732	Increased	225141_at	NFATC3	2.553	Increased	202556_s_at	MCRS1	2.165	Decreased
217371_s_at	IL15	2.953	Increased	218030_at	GIT1	3.032	Increased	203395_s_at	HES1	2.898	Decreased
205887_x_at	MSH3	2.1	Increased	218750_at	TAF1D	2.304	Increased	203184_at	FBN2	2.145	Decreased
1568865_at	FNTB	2.556	Increased	201795_at	LBR	2.383	Increased	203904_x_at	CD82	2.382	Decreased
208296_x_at	TNFAIP8	2.118	Increased	210045_at	IDH2	2.018	Increased	206558_at	SIM2	2.745	Decreased
226534_at	KITLG	2.22	Increased	1553810_a_at	KIAA1524	2.457	Increased	203665_at	HMOX1	2.358	Decreased
214981_at	POSTN	2.095	Increased	209160_at	AKR1C3	2.281	Increased	203543_s_at	KLF9	2.577	Decreased
206026_s_at	TNFAIP6	2.008	Increased	232424_at	PRDM16	2.334	Increased	1557729_at	GRK5	2.098	Decreased
210135_s_at	SHOX2	2.194	Increased	1554509_a_at	FAM188A	-2.494	Increased	236028_at	IBSP	-3.459	Decreased
204457_s_at	GAS1	2.134	Increased	217494_s_at	PTENP1	-2.849	Increased	221539_at	EIF4EBP1	2.487	Decreased
221884_at	MECOM	2.004	Increased	216205_s_at	MFN2	-2.34	Increased	202430_s_at	PLSCR1	2.324	Decreased
209919_x_at	GGT1	2.63	Increased	200644_at	MARCKSL1	3.121	Increased	204054_at	PTEN	2.145	Decreased
229468_at	CDK3	2.022	Increased	217991_x_at	SSBP3	3.21	Increased	209617_s_at	CTNND2	-2.01	Decreased
1563182_at	ACVR1C	-2.241	Increased	211653_x_at	AKR1C1/AKR1C2	2.791	Increased	231697_s_at	VMP1	2.741	Decreased
227404_s_at	EGR1	2.543	Increased	228302_x_at	CAMK2N1	-3.3	Increased	210143_at	ANXA10	2.124	Decreased
210933_s_at	FSCN1	2.227	Increased	234040_at	HELLS	2.476	Increased	206233_at	B4GALT6	-2.767	Decreased
205357_s_at	AGTR1	2.987	Increased	218413_s_at	ZNF639	2.019	Increased	209705_at	MTF2	2.504	Decreased
209925_at	OCLN	-2.024	Increased	242907_at	GBP2	2.488	Increased	209802_at	PHLDA2	2.17	Decreased
205732_s_at	NCOA2	2.484	Increased	219377_at	GAREM	2.237	Increased	219685_at	TMEM35	-2.404	Decreased
215404_x_at	FGFR1	2.421	Increased	212865_s_at	COL14A1	2.516	Increased	219047_s_at	ZNF668	2.244	Decreased
235521_at	HOXA3	2.05	Increased	203572_s_at	TAF6	2.79	Decreased	64474_g_at	DGCR8	2.201	Decreased
218807_at	VAV3	2.555	Increased	232231_at	RUNX2	2.12	Decreased	205159_at	CSF2RB	-4.605	Decreased
202202_s_at	LAMA4	3.771	Increased	202931_x_at	BIN1	2.056	Decreased	236012_at	PSMF1	2.172	Affected

205500_at	C5	2.375	Increased	214433_s_at	SELENBP1	2.103	Decreased	235833_at	PPAT	3.266	Affected
204128_s_at	RFC3	2.06	Increased	228766_at	CD36	-4.325	Decreased	221261_x_at	MAGED4/MAGED4B	3.294	Affected
200951_s_at	CCND2	18.529	Increased	235300_x_at	RCHY1	2.569	Decreased	204983_s_at	GPC4	-2.902	Affected
222073_at	COL4A3	-2.278	Increased	222999_s_at	CCNL2	2.019	Decreased	231837_at	USP28	2.365	Affected
226731_at	ITGA1	2.816	Increased	208791_at	CLU	-2.492	Decreased	1554606_at	CEP120	2.141	Affected
227125_at	IFNAR2	2.053	Increased	203973_s_at	CEBPD	2.099	Decreased	224325_at	FZD8	2.488	Affected
222036_s_at	MCM4	2.31	Increased	209383_at	DDIT3	2.065	Decreased	217650_x_at	ST3GAL2	2.003	Affected
204061_at	PRKX	2.226	Increased	219266_at	ZNF350	2.143	Decreased	224022_x_at	WNT16	-3.268	Affected
228962_at	PDE4D	2.562	Increased	206825_at	OXTR	-2.396	Decreased	208962_s_at	FADS1	3.311	Affected
212672_at	ATM	2.024	Increased	205027_s_at	MAP3K8	2.11	Decreased	202948_at	IL1R1	2.512	Affected
225572_at	CREB1	2.243	Increased	205891_at	ADORA2B	2.166	Decreased	50277_at	GGA1	2.292	Affected
203046_s_at	TIMELESS	2.156	Increased	231947_at	MYCT1	-2.144	Decreased	1555843_at	HNRNPM	2.611	Affected
213943_at	TWIST1	2.021	Increased	212430_at	RBM38	2.254	Decreased	214157_at	GNAS	-2.307	Affected
209465_x_at	PTN	2.357	Increased	204159_at	CDKN2C	2.004	Decreased	215987_at	RAPGEF2	2.046	Affected
213506_at	F2RL1	4.829	Increased	212401_s_at	CDK11A/CDK11B	2.086	Decreased	216237_s_at	MCM5	2.753	Affected
231559_at	NNMT	2.208	Increased	205080_at	RARB	2.058	Decreased	206086_x_at	HFE	2.218	Affected
225740_x_at	MDM4	2.283	Increased	214727_at	BRCA2	2.489	Decreased	229807_s_at	MAZ	2	Affected
226636_at	PLD1	2.022	Increased	202718_at	IGFBP2	-2.099	Decreased	231002_s_at	RABEP1	2.266	Affected
227048_at	LAMA1	2.296	Increased	213811_x_at	TCF3	2.395	Decreased	209753_s_at	TMPO	2.116	Affected
230462_at	NUMB	2.763	Increased	202526_at	SMAD4	2.16	Decreased	201627_s_at	INSIG1	3.156	Affected
201727_s_at	ELAVL1	2.216	Increased	1567013_at	NFE2L2	3.005	Decreased	218019_s_at	PDXK	2.152	Affected
205204_at	NMB	2.09	Increased	234339_s_at	GLTSCR2	2.573	Decreased	204639_at	ADA	2.051	Affected
232044_at	RBBP6	2.324	Increased	206332_s_at	IFI16	2.24	Decreased	236223_s_at	RIT1	2.823	Affected
205394_at	CHEK1	2.071	Increased	228967_at	EIF1	2.161	Decreased	208913_at	GGA2	2.062	Affected
57532_at	DVL2	2.03	Increased	235593_at	ZEB2	2.106	Decreased	201286_at	SDC1	2.164	Affected
208937_s_at	ID1	2.256	Increased	1556583_a_at	SLC8A1	-2.156	Decreased	201106_at	GPX4	2.087	Affected

'Increased' means the genes up- or down-expression is predicted to promote proliferation of cells. 'Decreased' means the up- or down-expression is predicted to inhibit. 'Affected' means IPA could not predict whether the expression change promote or inhibit.

Table 3: 'Proliferation of cells' genes differentially expressed by at least 2-fold between hMSCs/CyclinD2 and hMSCs/Empty.

ID	Genes in dataset	Fold Change	Prediction (based on expression direction)
206649_s_at	TFE3	2.069	Increased
209960_at	HGF	2.732	Increased
217371_s_at	IL15	2.953	Increased
232231_at	RUNX2	2.12	Increased
235300_x_at	RCHY1	2.569	Increased
226534_at	KITLG	2.22	Increased
235423_at	ORC2	3.145	Increased
203973_s_at	CEBPD	2.099	Increased
227404_s_at	EGR1	2.543	Increased
205027_s_at	MAP3K8	2.11	Increased
212401_s_at	CDK11A/CDK11B	2.086	Increased
213811_x_at	TCF3	2.395	Increased
206332_s_at	IFI16	2.24	Increased
212672_at	ATM	2.024	Increased
235593_at	ZEB2	2.106	Increased
205394_at	CHEK1	2.071	Increased
208937_s_at	ID1	2.256	Increased
223570_at	MCM10	2.545	Increased
242979_at	IRS1	2.329	Increased
212148_at	PBX1	2.514	Decreased
204159_at	CDKN2C	2.004	Decreased
1552721_a_a	FGF1	-2.054	Decreased
205016_at	TGFA	-3.032	Decreased
201727_s_at	ELAVL1	2.216	Decreased
203395_s_at	HES1	2.898	Decreased

203665_at	HMOX1	2.358	Decreased
204054_at	PTEN	2.145	Decreased
219312_s_at	ZBTB10	2.171	Affected
229861_at	RFFL	2.017	Affected
228302_x_at	CAMK2N1	-3.3	Affected
235764_at	PRDM5	2.205	Affected
1554509_a_a	FAM188A	-2.494	Affected
232424_at	PRDM16	2.334	Affected
221539_at	EIF4EBP1	2.487	Affected
204986_s_at	TAOK2	-2.03	Affected
221577_x_at	GDF15	2.143	Affected
225740_x_at	MDM4	2.283	Affected
203046_s_at	TIMELESS	2.156	Affected
200951_s_at	CCND2	18.529	Affected
1567013_at	NFE2L2	3.005	Affected
215404_x_at	FGFR1	2.421	Affected
209753_s_at	TMPO	2.116	Affected
209383_at	DDIT3	2.065	Affected
229468_at	CDK3	2.022	Affected
204457_s_at	GAS1	2.134	Affected
218833_at	ZAK	2.125	Affected
214048_at	MBD4	2.383	Affected
205345_at	BARD1	2.029	Affected
206693_at	IL7	2.399	Affected
209292_at	ID4	3.609	Affected

'Increased' means the genes up- or down-expression is predicted to increase the function of interphase. 'Decreased' means the up- or down-expression is predicted to decrease. 'Affected' means IPA could not predict whether the expression change increase or decrease

Table 4. 'Interphase' genes differentially expressed by at least 2-fold between hMSCs/CyclinD2 and hMSCs/Empty.

cells, in which Rb is inactivated by human papillomavirus E6 and E7 proteins [24], did not promote cell proliferation (data not shown). On the other hand, increased expression of Cyclin D2 inhibits proliferation of primary human fibroblasts [25], indicating that Cyclin D2 has both positive and negative roles in the cell cycle, depending on cell type. We found that Cyclin D2 in hMSCs has a positive role in the cell cycle (Figure 2e).

TGF-β2 expression was suppressed in hMSCs/CyclinD2 compared with hMSC/Empty during culture (Figure 3a). We previously demonstrated that hMSC growth is reduced and TGF-β2 expression increases during long-term culture [26]. We also reported that fibroblast growth factor-2 (FGF-2) stimulates hMSC growth by suppressing the up-regulation of TGF-β2 [27]. It is unclear how overexpression of Cyclin D2 suppresses the TGF-β2 increase, but this suppression may be involved in the acceleration of hMSCs/CyclinD2. In contrast, the expression of p16, which is up-regulated with aging [28], was increased in both cell types during culture, indicating that not only hMSCs/Empty but also hMSCs/CyclinD2 were aging normally. The rate of increase in hMSCs/Cyclin D2 was higher than in hMSCs/Empty (Figure 3b), suggesting that the promotion of cell proliferation in hMSCs/CyclinD2 induced cellular senescence and enhanced p16 expression. p16 is a tumor suppressor gene [29]; thus, this increase in p16 expression probably prevented unlimited proliferation. Consistent with this notion, some Ewing's sarcomas contain a homozygous deletion of the p16 locus [16], possibly facilitating subsequent transformation.

In this study, overexpression of Cyclin D2 promoted proliferation of hMSCs but did not lead to unlimited proliferation. Other factors are required for the unlimited proliferation of hMSCs. IGF2BP1 was aberrantly expressed in Ewing's sarcoma (Table 1), consistent with a previous report of an association between increased IGF2BP1

expression and tumor progression in patients with lung cancer [30]. Thus, we attempted to transduce the IGF2BP1 gene into hMSCs, but IGF2BP1 expression was up-regulated by only 2-fold and transduction efficiency was low (data not shown). The cause for this inefficiency is unclear. Because the growth kinetics of IGF2BP1-transformed *E. coli* is quite slow (data not shown), it is likely that overexpression of IGF2BP1 is deleterious for hMSCs.

We did not tested whether the other genes listed in Table 1 affect proliferation of hMSCs, because these genes were not thought to directly affect the proliferation. Furthermore, not all Ewing's sarcomas express EWS-FLI-1: indeed, EWS-FLI-1 mRNA was not detected in Hs 822.T and Hs 863.T (data not shown). Thus, we did not transduce the EWS-FLI-1 gene into hMSCs. However, it is possible that the cooperation of these proteins is important for the development of Ewing's sarcoma. Thus, it would be interesting to transduce these genes into hMSCs in addition to Cyclin D2.

Conclusion

Cyclin D2 promotes hMSC proliferation and is a candidate biomaker for hMSC transformation.

Acknowledgements

References

1. Pittenger MF, Mackay AM, Beck SC, Jaiswal RK, Douglas R, et al. (1999) Multilineage potential of adult human mesenchymal stem cells. Science 284: 143-147.

2. Caplan AI, Bruder SP (2001) Mesenchymal stem cells: building blocks for molecular medicine in the 21st century. Trends Mol Med 7: 259-264.

3. Gojo S, Gojo N, Takeda Y, Mori T, Abe H, et al. (2003) In vivo cardiovasculogenesis by direct injection of isolated adult mesenchymal stem cells. Exp Cell Res 288: 51-59.

4. Wakitani S, Saito T, Caplan AI (1995) Myogenic cells derived from rat bone marrow mesenchymal stem cells exposed to 5-azacytidine. Muscle Nerve 18: 1417-1426.

5. Prockop DJ (1997) Marrow stromal cells as stem cells for nonhematopoietic tissues. Science 276: 71-74.

6. Ohgushi H, Caplan AI (1999) Stem cell technology and bioceramics: from cell to gene engineering. J Biomed Mater Res 48: 913-927.

7. Petite H, Viateau V, Bensaïd W, Meunier A, de Pollak C, et al. (2000) Tissue-engineered bone regeneration. Nat Biotechnol 18: 959-963.

8. Ochi M, Adachi N, Nobuto H, Yanada S, Ito Y, et al. (2004) Articular cartilage repair using tissue engineering technique-novel approach with minimally invasive procedure. Artif Organs 28: 28-32.

9. Riggi N, Cironi L, Provero P, Suvà ML, Kaloulis K, et al. (2005) Development of Ewing's sarcoma from primary bone marrow-derived mesenchymal progenitor cells. Cancer Res 65: 11459-11468.

10. Riggi N, Suvà ML, Suvà D, Cironi L, Provero P, et al. (2008) EWS-FLI-1 expression triggers a Ewing's sarcoma initiation program in primary human mesenchymal stem cells. Cancer Res 68: 2176-2185.

11. Miyagawa Y, Okita H, Nakaijima H, Horiuchi Y, Sato B, et al. (2008) Inducible expression of chimeric EWS/ETS proteins confers Ewing's family tumor-like phenotypes to human mesenchymal progenitor cells. Mol Cell Biol 28: 2125-2137.

12. Castillero-Trejo Y, Eliazer S, Xiang L, Richardson JA, Ilaria RL Jr (2005) Expression of the EWS/FLI-1 oncogene in murine primary bone-derived cells Results in EWS/FLI-1-dependent, ewing sarcoma-like tumors. Cancer Res 65: 8698-8705.

13. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, et al. (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature 359: 162-165.

14. Toretsky JA, Connell Y, Neckers L, Bhat NK (1997) Inhibition of EWS-FLI-1 fusion protein with antisense oligodeoxynucleotides. J Neurooncol 31: 9-16.

15. Tanaka K, Iwakuma T, Harimaya K, Sato H, Iwamoto Y (1997) EWS-Fli1 antisense oligodeoxynucleotide inhibits proliferation of human Ewing's sarcoma and primitive neuroectodermal tumor cells. J Clin Invest 99: 239-247.

16. Deneen B, Denny CT (2001) Loss of p16 pathways stabilizes EWS/FLI1 expression and complements EWS/FLI1 mediated transformation. Oncogene 20: 6731-6741.

17. Lessnick SL, Dacwag CS, Golub TR (2002) The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts. Cancer Cell 1: 393-401.

18. Takano Y, Kato Y, Masuda M, Ohshima Y, Okayasu I (1999) Cyclin D2, but not cyclin D1, overexpression closely correlates with gastric cancer progression and prognosis. J Pathol 189: 194-200.

19. Takano Y, Kato Y, van Diest PJ, Masuda M, Mitomi H, et al. (2000) Cyclin D2 overexpression and lack of p27 correlate positively and cyclin E inversely with a poor prognosis in gastric cancer cases. Am J Pathol 156: 585-594.

20. Mermelshtein A, Gerson A, Walfisch S, Delgado B, Shechter-Maor G, et al. (2005) Expression of D-type cyclins in colon cancer and in cell lines from colon carcinomas. Br J Cancer 93: 338-345.

21. Dhillon VS, Shahid M, Husain SA (2004) CpG methylation of the FHIT, FANCF, cyclin-D2, BRCA2 and RUNX3 genes in Granulosa cell tumors (GCTs) of ovarian origin. Mol Cancer 3: 33.

22. Sherr CJ (1993) Mammalian G1 cyclins. Cell 73: 1059-1065.

23. Bartkova J, Rajpert-de Meyts E, Skakkebaek NE, Bartek J (1999) D-type cyclins in adult human testis and testicular cancer: relation to cell type, proliferation, differentiation, and malignancy. J Pathol 187: 573-581.

24. Goodwin EC, DiMaio D (2000) Repression of human papillomavirus oncogenes in HeLa cervical carcinoma cells causes the orderly reactivation of dormant tumor suppressor pathways. Proc Natl Acad Sci U S A 97: 12513-12518.

25. Meyyappan M, Wong H, Hull C, Riabowol KT (1998) Increased expression of cyclin D2 during multiple states of growth arrest in primary and established cells. Mol Cell Biol 18: 3163-3172.

26. Sawada R, Ito T, Tsuchiya T (2006) Changes in expression of genes related to cell proliferation in human mesenchymal stem cells during in vitro culture in comparison with cancer cells. J Artif Organs 9: 179-184.

27. Ito T, Sawada R, Fujiwara Y, Seyama Y, Tsuchiya T (2007) FGF-2 suppresses cellular senescence of human mesenchymal stem cells by down-regulation of TGF-beta2. Biochem Biophys Res Commun 359: 108-114.

28. Krishnamurthy J, Ramsey MR, Ligon KL, Torrice C, Koh A, et al. (2006) p16INK4a induces an age-dependent decline in islet regenerative potential. Nature 443: 453-457.

29. Liggett WH Jr, Sidransky D (1998) Role of the p16 tumor suppressor gene in cancer. J Clin Oncol 16: 1197-1206.

30. Kato T, Hayama S, Yamabuki T, Ishikawa N, Miyamoto M, et al. (2007) Increased expression of insulin-like growth factor-II messenger RNA-binding protein 1 is associated with tumor progression in patients with lung cancer. Clin Cancer Res 13: 434-442.

Short paper

Characterization of the cell growth analysis for detection of immortal cellular impurities in human mesenchymal stem cells

CrossMark

Ken Kono [a, 1], Nozomi Takada [b, c, 1], Satoshi Yasuda [b, d], Rumi Sawada [a], Shingo Niimi [a], Akifumi Matsuyama [c], Yoji Sato [b, d, e, f, g, *]

[a] Division of Medical Devices, National Institute of Health Sciences, 1-18-1 Kami-yoga, Setagaya, Tokyo 158-8501, Japan
[b] Division of Cellular & Gene Therapy Products, National Institute of Health Sciences, 1-18-1 Kami-yoga, Setagaya, Tokyo 158-8501, Japan
[c] Research on Disease Bioresources, Platform of Therapeutics for Rare Disease and Health Policy, National Institute of Biomedical Innovation, Kobe International Business Center Rm#602, 5-5-2 Minatojima-Minami-Machi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
[d] Foundation for Biomedical Research and Innovation, Hyogo, Japan
[e] Department of Quality Assurance Science for Pharmaceuticals, Graduate School of Pharmaceutical Sciences, Nagoya City University, Aichi, Japan
[f] Department of Cellular & Gene Therapy Products, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan
[g] Department of Translational Pharmaceutical Sciences, Graduated School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan

ABSTRACT

The analysis of in vitro cell senescence/growth after serial passaging can be one of ways to show the absence of immortalized cells, which are frequently tumorigenic, in human cell-processed therapeutic products (hCTPs). However, the performance of the cell growth analysis for detection of the immortalized cellular impurities has never been evaluated. In the present study, we examined the growth rates of human mesenchymal stem cells (hMSCs, passage 5 ($P = 5$)) contaminated with various doses of HeLa cells, and compared with that of hMSCs alone. The growth rates of the contaminated hMSCs were comparable to that of hMSCs alone at $P = 5$, but significantly increased at $P = 6$ (0.1% and 0.01% HeLa) or $P = 7$ (0.001% HeLa) within 30 days. These findings suggest that the cell growth analysis is a simple and sensitive method to detect immortalized cellular impurities in hCTPs derived from human somatic cells.
© 2014 The Authors. Published by Elsevier Ltd on behalf of The International Alliance for Biological Standardization. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction

Human cell-processed therapeutic products (hCTPs) are expected to provide novel breakthrough therapies for currently life-threatening or incurable diseases. In the clinical applications of hCTPs to patients, however, one of the major concerns is the tumorigenic cellular impurities in the products. Since pluripotent stem cells (PSCs), such as embryonic stem cells and induced pluripotent stem cells, are tumorigenic [1–3], there is a risk of tumor formation if the products contain the residual undifferentiated PSCs [4]. On the other hand, somatic cells are considered to have little tumorigenic potential even after substantial manipulations like in vitro expansion, because they consistently pass into senescence [5]. Malignant transformation of the cells is believed to occur through multiple processes involving the accumulation of mutations in key regulatory genes that promote cell survival and proliferation [6,7]. Although a few individual groups reported the spontaneous transformation of human mesenchymal stem cells (hMSCs) during in vitro culture [8–11], two of them retracted their papers because the results appeared to be attributable to contamination with tumorigenic cells (fibrosarcoma, osteosarcoma, or glioma cell lines) [12,13]. The rest of the groups found the immortalization of the cells, which is closely associated with tumorigenicity, during in vitro culture, indicating that the good practices to avoid contamination with tumorigenic cells and the monitoring of cell growth are critical for the quality control of hCTPs derived from human somatic cells.

* Corresponding author. Division of Cellular and Gene Therapy Products, National Institute of Health Sciences, 1-18-1 Kami-yoga, Setagaya, Tokyo 158-8501, Japan.
E-mail addresses: kenkono@nihs.go.jp (K. Kono), takada.n@nihs.go.jp (N. Takada), yasuda@nihs.go.jp (S. Yasuda), rsawada@nihs.go.jp (R. Sawada), niimi@nihs.go.jp (S. Niimi), akifumi-matsuyama@umin.ac.jp (A. Matsuyama), yoji@nihs.go.jp (Y. Sato).
[1] These authors contributed equally to this work.

Abbreviations

hCTP	human cell-processed therapeutic product
PSC	pluripotent stem cell
hMSC	human mesenchymal stem cell
STR	short tandem repeat
$P = n$	passage n
PBS	phosphate buffered saline
RT	room temperature
HPV	human papillomavirus

Cross contamination of cells with unidentified cells is usually evaluated by the short tandem repeat (STR) analysis [14]. However, the cell growth analysis which simply monitors the cell proliferation for a limited period may be adequately sensitive for the detection of the contamination of somatic cells with immortalized/tumorigenic cells, because somatic cells usually show slower growth, compared with that of immortalized/tumorigenic cells, as well as the attenuation of the growth after serial passaging [15–17]. In fact, the European Medicines Agency has considered that the evaluation of *in vitro* cell senescence after serial passaging is sufficient to prove the absence of immortalized/tumorigenic cells in a somatic cell-based product [18]. However, the performance of the cell growth analysis for detection of the immortalized/tumorigenic cellular impurities in somatic cells has never been studied. In the present study, we examined the growth of hMSCs contaminated with various doses of HeLa cells, a well-known cancer cell line, to determine the sensitivity of the cell growth analysis for the detection of the immortalized/tumorigenic cells in human somatic cells.

2. Materials and methods

2.1. Cells

hMSCs (Lonza, Walkersville, MD) at passage 2 ($P = 2$) were cultured in MSCGM BulletKit, a mesenchymal stem cell basal medium with mesenchymal cell growth supplement, L-glutamine, and gentamycin/amphotericin-B (Lonza). HeLa cells (the Health Science Research Resources Bank, Osaka, Japan) were maintained in Eagle's minimum essential medium (Sigma), supplemented with 10% fetal bovine serum (FBS; Sigma), 0.1 mM non-essential amino acids (Life Technologies), 50 U/ml penicillin, and 50 µg/ml streptomycin (Life Technologies). Cells were cultured in a humidified atmosphere of 5% CO_2 and 95% air at 37 °C, and were passaged upon reaching 90% confluence.

2.2. Cell growth analysis

At $P = 5$ of hMSCs, 1×10^6 of hMSCs were mixed with 1000, 100, or 10 of HeLa cells and seeded into T175 flasks (Corning). The cells were maintained in 40 ml of Dulbecco's Modified Eagle's medium (DMEM; Gibco) supplemented with 10% FBS, 50 U/ml penicillin, and 50 µg/ml streptomycin. Upon reaching approximately 90% confluence, the cells were washed with phosphate buffered saline (PBS) and treated with 0.05% trypsin-EDTA solution (Gibco) for detachment from the flasks. The cells were centrifuged at 450 × g for 5 min and suspended with the fresh culture medium. Aliquots of the suspended cells were stained with Trypan Blue solution and counted by Countess Automated Cell Counter (Invitrogen) according to the manufacture's protocol. One million cells in the suspension were re-seeded into T175 flasks and cultured until the next passage. This process was repeated by $P = 10$. The growth rate (R_n) at $P = n$ was calculated by the following equation:

$$R_n = [\log_2(N_{n+1} - N_n)]/(D_{n+1} - D_n)$$

where N_k and D_k are the number of accumulated cells and the date at $P = k$, respectively.

2.3. Immunofluorescence microscopy

hMSCs contaminated with HeLa cells were fixed with 4% paraformaldehyde in PBS (Nacalai Tesque) for 10 min at room temperature (RT) and blocked in Blocking One (Nacalai Tesque) for 30 min at RT. The cells, then, were incubated with anti-HPV18 E7 antibody (8E2) (abcam) diluted at 1:500 in the blocking solution (PBS containing 5% Blocking One) for 1 h at RT for primary staining, and secondarily stained with goat anti-mouse IgG Alexa Fluor 488 (1:1000; Invitrogen) in the blocking solution for 45 min at RT. The cells were mounted with VECTASGIELD mounting medium with DAPI (VECTOR) and observed with a fluorescence microscope (IX71, Olympus).

3. Results and discussion

In the present study, we added 1000, 100, or 10 of HeLa cells to 1×10^6 of hMSCs of passage 5 ($P = 5$) and compared their growth with that of hMSCs alone (HeLa 0) until $P = 10$. The growth curves of three lots of hMSCs and the contaminated hMSCs are shown in Fig. 1. The cell numbers of HeLa 0 and the contaminated hMSCs were comparable at $P = 5$. The growth of HeLa 0 was constant during the early culture and getting slower with time (Fig. 1), while the growth of the contaminated cells was accelerated.

To confirm that the increases in the growth were attributable to the contamination with HeLa cells, we observed the cells with phase contrast microscopy. In the images of the contaminated hMSCs, we found small cells clearly different from hMSCs (Fig. 2A), and their relative abundance increased every passage. Because HeLa cells are infected with human papillomaviruses (HPV), we performed immunofluorescence analysis using HPV18 E7 antibody and confirmed that the cells were HeLa cells not transformed hMSCs (Fig. 2B). At $P = 10$, hMSCs were hardly identified in images of HeLa 1000 (Fig. 2C), because almost all of hMSCs were exchanged for HeLa cells at the five passages.

Next, we examined the growth rates of the contaminated cells (Fig. 3A). They were comparable to that of HeLa 0 at $P = 5$, and got significantly increased at $P = 6$ (HeLa 1000 and HeLa 100) or $P = 7$ (HeLa 10). These results indicated that the gross proliferation rate was not influenced by the spiked cells at $P = 5$ and then the population of HeLa cells in hMSCs increased in dose- and time-dependent manner. Eventually, the growth rate (doubling/day) of the contaminated hMSCs increased, and then reached plateau. The average growth rate of the contaminated cells at $P = 9$ and 10 was 0.73, suggesting that the growth rate of HeLa cells was approximately 0.7 in this culture condition. The average growth rates of the three lots are plotted along the passage number in Fig. 3B. The growth rates of HeLa 1000 and HeLa 100 at $P = 7$ and HeLa 10 at $P = 8$ were significantly increased compared with the growth rates at $P = 5$ (*$P < 0.05$, two-way repeated measures ANOVA and Student-Newman–Keulis test). These results indicate that the cell cultures longer than $P = 7$ (about 20 days) and $P = 8$ (about 30 days) detect cross-contaminations of 100 (0.01%) and 10 (0.001%) HeLa cells, respectively, assuming that 10^6 hMSCs were contaminated at $P = 5$.
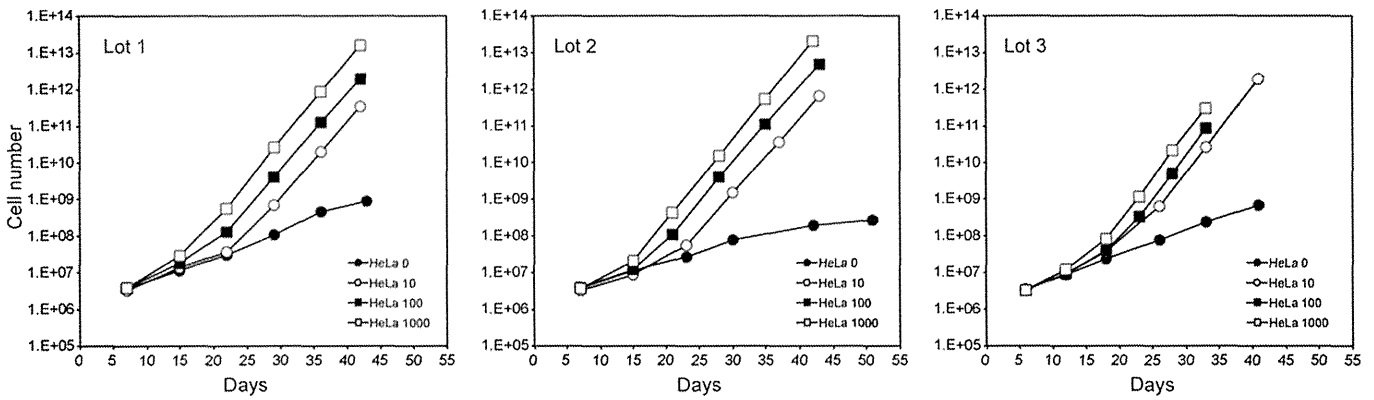
Fig. 1. Cell growth analysis of hMSCs contaminated with HeLa cells. At the passage 5 ($P = 5$) of hMSCs, 1×10^6 of hMSCs were mixed with 0, 10, 100, or 1000 of HeLa cells (HeLa 0, HeLa 10, HeLa 100, or HeLa 1000). Cells were passaged and counted at the indicated day from $P = 5$. The results of three lots of hMSCs are presented (Lot 1, 2, and 3).

The tumorigenicity has been evaluated by *in vitro* assays (e.g., soft agar formation assay, karyotype analysis) and/or *in vivo* assays (transplantation into immunodeficient animals) [8,9,19,20]. When these assays are performed to detect a trace amount of tumorigenic cells in hCTPs, they need to be of high sensitivity. Recently, we developed a highly sensitive *in vivo* tumorigenicity test using severely immunocompromised mice, NOG mice, in combination with Matrigel. Subcutaneous transplantation into NOG mice with Matrigel allowed inoculation with 1×10^7 cells and actually achieved detection of 0.002% HeLa cells spiked into hMSCs in a half of the mice (unpublished data). However, *in vivo* tumorigenicity test using immunodeficient animals requires a specific facility, and takes 3–4 months. In contrast, the cell growth analysis is not only simple and economical, but also detects as few as 0.001% immortal cellular impurities in hMSCs within 30 days, in case that the growth properties of the immortal cells are comparable to those of HeLa cells. Although immortality does not necessarily indicate tumorigenicity, it is known to be closely associated with cell transformation in many cases of tumorigenesis. Therefore, the present study can be said to be the first scientific basis for the usefulness of the cell growth analysis as one of tumorigenicity tests for hCTPs derived from human somatic cells.
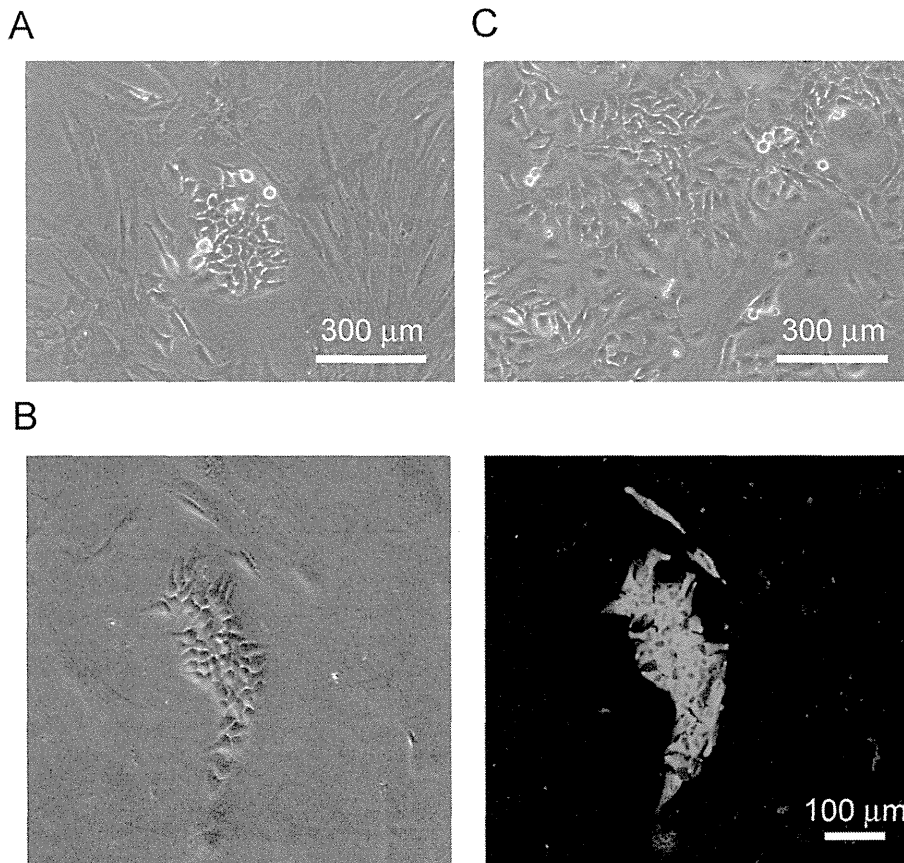

Fig. 2. Phase contrast and immunofluorescence microscopy of hMSCs contaminated with HeLa cells. (A) Representative image of the contaminated cells (HeLa 100 at $P = 8$) is presented. Red arrow pointed at the small size cells which were clearly different from hMSCs. (B) Phase contrast and immunofluorescence microscopy of HeLa 1000 at day 5 was conducted by using anti-HPV E7 antibody. Green: Alexa Fluor 488 goat anti-mouse IgG; blue: DAPI. (C) Representative image of HeLa 1000 at P10 is presented.
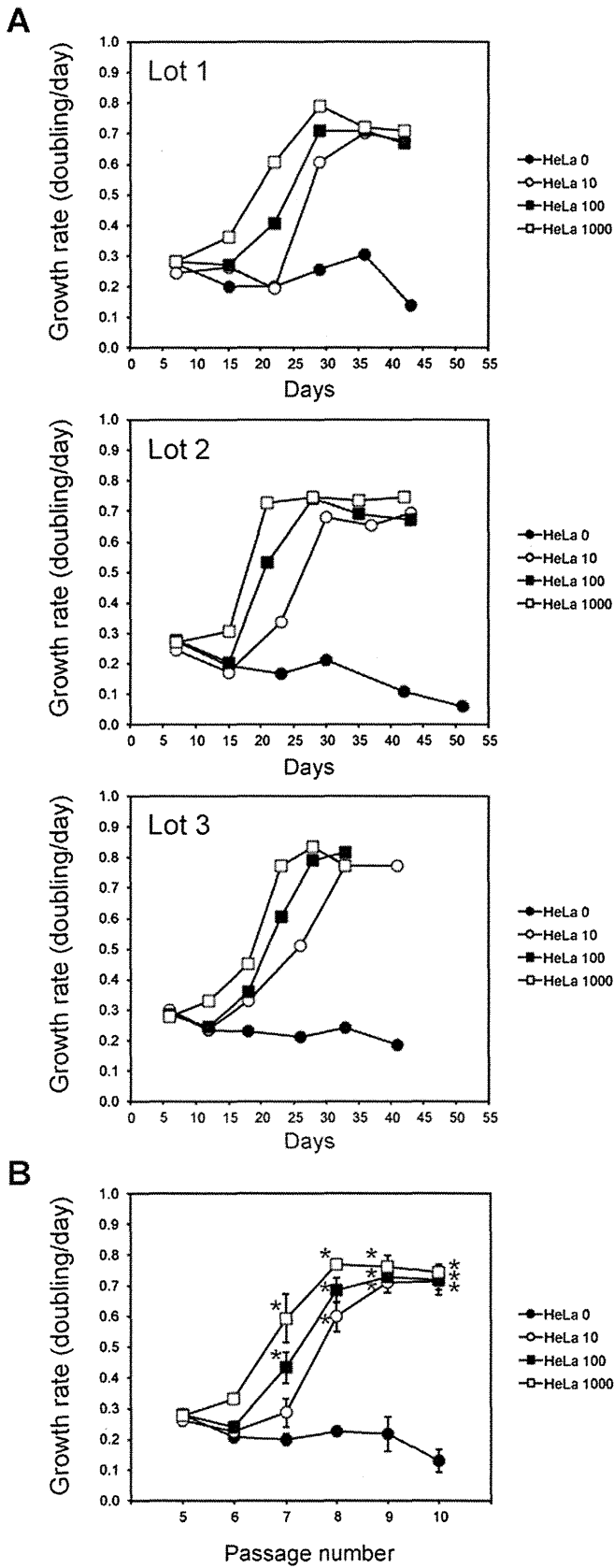
Fig. 3. Cell growth rate analysis of hMSCs contaminated with HeLa cells. (A) The Y-axis indicates the growth rate (doubling per day) of hMSCs. The results of three lots of hMSCs are presented (Lot 1, 2, and 3). (B) The growth rates of three lots of hMSCs are plotted along the passage numbers. Data are represented as mean ± S.E.M. of the three lots. Statistical significance was determined using two-way repeated measures ANOVA and Student-Newman–Keulis's post-hoc test (*$P < 0.05$ compared with the rate at $P = 5$).

Author contributions

Conceived and designed the experiments: KK NT SY RS YS. Performed the experiments: KK NT. Analyzed and Interpreted the data: KK NT SY RS SN AM YS. Contributed regents/materials/analysis tools: SN AM YS. Wrote the paper: KK NT SY YS. Acquired the funding: AM YS.

Competing interests

The authors have declared that no competing interests exist.

Acknowledgments

References

[1] Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861–72.

[2] Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145–7.

[3] Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007;318:1917–20.

[4] Kuroda T, Yasuda S, Sato Y. Tumorigenicity studies for human pluripotent stem cell-derived products. Biol Pharm Bull 2013;36:189–92.

[5] Prockop DJ. Defining the probability that a cell therapy will produce a malignancy. Mol Ther 2010;18:1249–50.

[6] Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med 2004;10:789–99.

[7] Blagosklonny MV. Target for cancer therapy: proliferating cells or stem cells. Leukemia 2006;20:385–91.

[8] Rubio D, Garcia-Castro J, Martin MC, de la Fuente R, Cigudosa JC, Lloyd AC, et al. Spontaneous human adult stem cell transformation. Cancer Res 2005;65:3035–9.

[9] Rosland GV, Svendsen A, Torsvik A, Sobala E, McCormack E, Immervoll H, et al. Long-term cultures of bone marrow-derived human mesenchymal stem cells frequently undergo spontaneous malignant transformation. Cancer Res 2009;69:5331–9.

[10] Wang Y, Huso DL, Harrington J, Kellner J, Jeong DK, Turney J, et al. Outgrowth of a transformed cell population derived from normal human BM mesenchymal stem cell culture. Cytotherapy 2005;7:509–19.

[11] Tang DQ, Wang Q, Burkhardt BR, Litherland SA, Atkinson MA, Yang LJ. In vitro generation of functional insulin-producing cells from human bone marrow-derived stem cells, but long-term culture running risk of malignant transformation. Am J Stem Cells 2012;1:114–27.

[12] Torsvik A, Rosland GV, Svendsen A, Molven A, Immervoll H, McCormack E, et al. Spontaneous malignant transformation of human mesenchymal stem cells reflects cross-contamination: putting the research field on track — letter. Cancer Res 2010;70:6393–6.

[13] de la Fuente R, Bernad A, Garcia-Castro J, Martin MC, Cigudosa JC. Retraction: spontaneous human adult stem cell transformation. Cancer Res 2010;70:6682.

[14] Dirks WG, MacLeod RA, Nakamura Y, Kohara A, Reid Y, Milch H, et al. Cell line cross-contamination initiative: an interactive reference database of STR profiles covering common cancer cell lines. Int J Cancer 2010;126:303–4.

[15] Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. Exp Cell Res 1961;25:585–621.

[16] Hayflick L. The cell biology of aging. Clin Geriatr Med 1985;1:15–27.

[17] Hayflick L. The limited in vitro lifetime of human diploid cell strains. Exp Cell Res 1965;37:614–36.

[18] European medicines Agency EMEA/724428/2009 assessment report for ChondroCelect. 2009.

[19] Kuroda T, Yasuda S, Kusakawa S, Hirata N, Kanda Y, Suzuki K, et al. Highly sensitive in vitro methods for detection of residual undifferentiated cells in retinal pigment epithelial cells derived from human iPS cells. PLoS One 2012;7:e37342.

[20] Rangarajan A, Hong SJ, Gifford A, Weinberg RA. Species- and cell type-specific requirements for cellular transformation. Cancer Cell 2004;6:171–83.

The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice

Atsushi Fukuda[1], Junko Tomikawa[2], Takumi Miura[1], Kenichiro Hata[2], Kazuhiko Nakabayashi[2], Kevin Eggan[3], Hidenori Akutsu[1] & Akihiro Umezawa[1]

Maintaining a single active X-chromosome by repressing *Xist* is crucial for embryonic development in mice. Although the *Xist* activator RNF12/RLIM is present as a maternal factor, maternal *Xist* (*Xm-Xist*) is repressed during preimplantation phases to establish imprinted X-chromosome inactivation (XCI). Here we show, using a highly reproducible chromatin immunoprecipitation method that facilitates chromatin analysis of preimplantation embryos, that H3K9me3 is enriched at the *Xist* promoter region, preventing Xm-*Xist* activation by RNF12. The high levels of H3K9me3 at the *Xist* promoter region are lost in embryonic stem (ES) cells, and ES-cloned embryos show RNF12-dependent *Xist* expression. Moreover, lack of Xm-XCI in the trophectoderm, rather than loss of paternally expressed imprinted genes, is the primary cause of embryonic lethality in 70–80% of parthenogenotes immediately after implantation. This study reveals that H3K9me3 is involved in the imprinting that silences Xm-*Xist*. Our findings highlight the role of maternal-specific H3K9me3 modification in embryo development.

[1] Department of Reproductive Biology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535, Japan. [2] Department of Maternal-Foetal Biology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya, Tokyo 157-8535, Japan. [3] The Howard Hughes Medical Institute, Harvard Stem Cell Institute and the Department of Stem Cell and Regenerative Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA. Correspondence and requests for materials should be addressed to H.A. (email: akutsu-h@ncchd.go.jp).

o maintain proper dosage compensation in mammals, one of the two X chromosomes in the female is inactivated[1,2]. In establishment of X-chromosome inactivation (XCI), a large non-coding RNA, *Xist*, is expressed and this non-coding RNA then covers the entire X chromosome in *cis*[1–3]. In mice, two types of XCI occur during female embryonic development. One type involves random XCI, which is observed in cells derived from epiblasts, and one of the two X chromosomes (paternal or maternal) is randomly inactivated. The other involves imprinted XCI (iXCI), which is observed in extra-embryonic tissues and causes XCI of the paternal X chromosome (Xp)[4]. The initiation of iXCI begins at early preimplantation in embryos and Xp-*Xist* is expressed around the four-cell stage[1]. A recent study showed that a maternal factor, the E3 ubiquitin ligase RNF12, is the primary factor responsible for Xp-*Xist* activation[5]. Interestingly, although RNF12 is abundant as a maternal factor in oocytes, Xm-*Xist* is not expressed. Moreover, maternal *Xist* (Xm-*Xist*)-specific imprints, which are refractory to the Xm-*Xist* activation induced by RNF12, are imposed during oogenesis[6]. *Xist* expression analysis using *de novo* DNA methyltransferase (*Dnmt3a/b*) maternal knockout mice demonstrated that *Xist* expression during preimplantation is independent of DNA methylation[7], implying that other epigenetic factors are associated with Xm-*Xist* silencing. However, the nature of these Xm-specific epigenetic modifications is unknown.

A gene-knockout study demonstrated that loss of Xp-*Xist* expression critically affects postimplantation female development due to lack of iXCI, which causes overexpression of X-linked genes in extra-embryonic tissues[8]. Similar to the phenotype observed in Xp-*Xist*-knockout mice, parthenogenetic embryos (PEs) composed of two X chromosomes show increased expression of X-linked genes, as compared with fertilized females, because of the low expression of *Xist*[9]. One of the interesting phenomena observed in PEs is the dramatic developmental failure that occurs immediately after implantation. Around 70–80% of embryos die before embryonic day (E) 9.5, which is the limit of development for PEs[10]. However, it is unknown whether the primary cause of rapid developmental failure in postimplantation PEs is the loss of iXCI or the loss of expression of autosomal paternally imprinted genes[11,12].

The global epigenetic asymmetry of parental genomes in zygotes is retained during early preimplantation phases in mice and changes in gene expression occur in discrete stages to confer totipotency[13,14]. Interestingly, transcriptionally repressive marks, such as histone H3 lysine 9 di-/trimethylation (H3K9me2/3), are specifically imposed on maternal genomes at the zygote stage[13]. Although the regulation of imprinted genes mostly depends on DNA methylation, some imprinted genes are regulated by these histone modifications[15,16]. Thus, Xm-*Xist* silencing machinery may be associated with histone modifications.

Here we reveal that silencing of Xm-*Xist* in preimplantation embryos involves modification of H3K9me3. By using a new chromatin immunoprecipitation (ChIP) method that facilitates chromatin analysis in preimplantation embryos, we show that the *Xist* promoter on the Xm is highly enriched for H3K9me3 at the four-cell stage. This enrichment is lost in the morula and in male embryonic stem (ES) cells. Furthermore, we demonstrate that early loss of H3K9me3 at the *Xist* promoter leads to precocious Xm-*Xist* activation in a Rnf12-dependent manner. Moreover, we demonstrate that establishment of Xm-XCI in the trophectoderm allows PEs to develop at the postimplantation stage without the expression of paternally imprinted genes on autosomes. Therefore, these data indicate that the primary cause of embryonic lethality immediately after implantation in most PEs is loss of XCI rather than loss of the expression of paternally imprinted genes located on autosomes. Our study revealed that silencing of Xm-*Xist* by imprinting to establish iXCI involves H3K9me3, and this finding is expected to resolve the longstanding issues that have limited our general understanding of XCI in mice.

Results

Changes in histone modifications cause Xm-*Xist* derepression. Histone repressive marks, such as H3K9me2/3 and H3K27me3, are specifically imposed on maternal genomes[13]. To investigate the role of maternal-specific modifications in imprinted *Xist* expression, we focused on *Kdm3a* and *Kdm4b*, which encode histone demethylases specific for H3K9me1/2 and H3K9me3 (refs 17,18), respectively. Reverse transcription–PCR analysis showed that oocytes express low levels of *Kdm3a* and *Kdm4b* (Supplementary Fig. 1). Immunofluorescence (IF) analyses revealed that zygotes injected with polyadenylated *Kdm3a* and *Kdm4b* messenger RNAs expressed significantly lower levels of maternal H3K9me2 and H3K9me3, respectively (Fig. 1a–d). Ectopic expression of *Kdm3a* and *Kdm4b* did not affect H3K9me3 or H3K9me2 marks, respectively (Supplementary Fig. 2). We reasoned that if Xm-specific modifications that prevent *Xist* activation were erased by these epigenetic modifiers, Xm-*Xist* would be expressed at the four-cell stage, which is when Xp-*Xist* expression commences.

To facilitate analysis of Xm-*Xist* expression, we used PEs (Fig. 1e). PEs possess two copies of Xm, and Xm-*Xist* is never expressed at the four-cell stage[19]. Xm-*Xist* expression in four-cell PEs, cultured for 48 h, was determined using quantitative real-time PCR (qPCR). Consistent with a previous report[19], Xm-*Xist* was not detectably expressed in most intact (not injected) PEs and PEs injected with *Egfp* mRNA (*Egfp*-PEs; Fig. 1f). Approximately 75% of PEs injected with *Kdm3a* mRNA (*Kdm3a*-PEs) did not detectably express *Xist*. However, Xm-*Xist* expression was detected in all PEs injected with *Kdm4b* mRNA (*Kdm4b*-PEs;

Figure 1 | Alterations in histone modifications derepress Xm-*Xist* expression. (**a–d**) Oocytes injected with *Kdm3a* (**a,b**), *Kdm4b* (**c,d**) or *Egfp* mRNAs were subjected to ICSI. After 7–8 h, embryos were fixed and analysed for H3K9me2 (**a**) and H3K9me3 (**c**) using IF. Nuclei stained with 4′,6-diamidino-2-phenylindole (DAPI) are shown in blue. Representative images are presented on the left. The box-and-whisker plot shows the ratio of maternal to paternal signal intensities. The horizontal line indicates the median. The *P*-values were calculated using the Mann–Whitney *U*-test (*U*-test). Pb, polar body; *n*, number of embryos analysed (**b,d**). (**e**) Schema of the generation of PEs with altered histone modifications. To examine the effects of histone demethylation on Xm-*Xist* expression, either H3K9me2 demethylase (*Kdm3a*) or H3K9me3 demethylase (*Kdm4b*) mRNAs were injected into MII oocytes that were then activated. To assess the effects of inhibition of histone deacetylation on Xm-*Xist* expression, oocytes were activated and incubated in the presence of TSA for 24 h. After 48 h, ten four-cell PEs were pooled and analysed as one biological replicate using qPCR. (**f**) Analysis of Xm-*Xist* expression at the four-cell stage. The expression level of Xm-*Xist* in female embryos derived from IVF was defined as 1. One or two asterisks indicate Xm-*Xist* expression in one or two replicates, respectively. The *P*-values were determined using Student's *t*-tests. Error bars indicate the mean ± s.e.m. (**g–i**) *Xist* FISH analysis of *Kdm4b*- and *Egfp* + TSA-PEs at the four-cell stage. (**g**) Representative images of FISH results. (**h**) Ratio of cells with *Xist* signal to the total number of cells. *n*, number of interphase cells analysed. (**i**) Ratio of cells with biallelic expression to total cells. The detailed FISH results are shown in Supplementary Table 1. Scale bars, 20 μm.

— 237 —