

RESEARCH ARTICLE

Open Access



Morphology-based noninvasive early prediction of serial-passage potency enhances the selection of clone-derived high-potency cell bank from mesenchymal stem cells

Takashi Suyama^{1,2†}, Yuto Takemoto^{3†}, Hiromi Miyachi², Yuko Kato^{1,2}, Yumi Matsuzaki^{1,2*} and Ryuji Kato^{3,4*} 

Abstract

Background: Rapidly expanding clones (RECs) are one of the single-cell-derived mesenchymal stem cell clones sorted from human bone marrow mononuclear cells (BMMCs), which possess advantageous features. The RECs exhibit long-lasting proliferation potency that allows more than 10 repeated serial passages in vitro, considerably benefiting the manufacturing process of allogenic MSC-based therapeutic products. Although RECs aid the preparation of large-variation clone libraries for a greedy selection of better-quality clones, such a selection is only possible by establishing multiple-candidate cell banks for quality comparisons. Thus, there is a high demand for a novel method that can predict “low-risk and high-potency clones” early and in a feasible manner given the excessive cost and effort required to maintain such an establishment.

Methods: LNGFR and Thy-1 co-positive cells from BMMCs were single-cell-sorted into 96-well plates, and only fast-growing clones that reached confluency in 2 weeks were picked up and passaged as RECs. Fifteen RECs were prepared as passage 3 (P3) cryostock as the primary cell bank. From this cryostock, RECs were passaged until their proliferation limitation; their serial-passage limitation numbers were labeled as serial-passage potencies. At the P1 stage, phase-contrast microscopic images were obtained over 6–90 h to identify time-course changes of 24 morphological descriptors describing cell population information. Machine learning models were constructed using the morphological descriptors for predicting serial-passage potencies. The time window and field-of-view-number effects were evaluated to identify the most efficient image data usage condition for realizing high-performance serial-passage potency models.

Results: Serial-passage test results indicated variations of 7–13-repeated serial-passage potencies within RECs. Such potency values were predicted quantitatively with high performance (RMSE < 1.0) from P1 morphological profiles

[†]Takashi Suyama and Yuto Takemoto contributed equally.

*Correspondence: matsuzak@med.shimane-u.ac.jp; kato-r@ps.nagoya-u.ac.jp

¹ Department of Life Science, Faculty of Medicine, Shimane University, 89-1 Enya-cho, Izumo, Shimane 693-8501, Japan

³ Department of Basic Medicinal Sciences, Graduate School of Pharmaceutical Sciences, Nagoya University, Tokai National Higher Education and Research System, Furocho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

Full list of author information is available at the end of the article



using a LASSO model. The earliest and minimum effort predictions require 6–30 h with 40 FOVs and 6–90 h with 15 FOVs, respectively.

Conclusion: We successfully developed a noninvasive morphology-based machine learning model to enhance the efficiency of establishing cell banks with single-cell-derived RECs for quantitatively predicting the future serial-passage potencies of clones. Conventional methods that can make noninvasive and quantitative predictions without wasting precious cells in the early stage are lacking; the proposed method will provide a more efficient and robust cell bank establishment process for allogenic therapeutic product manufacturing.

Keywords: Morphological analysis, Mesenchymal stem cells, LNGFR, THY-1, Rapidly expanding clone (REC), Cell bank establishment, Serial-passage potency, Prediction model

Background

Mesenchymal stem cells (MSCs) are the most widely studied stem cells for cell-based therapeutic applications [1–3]. It is known that a variety of cell types exist in MSCs [4–7] because conventional MSC processing simply collects the adherent cell fraction from a cell suspension mixture [8, 9]. LNGFR (CD271) and THY-1 (CD90) co-positive cells (LT cells) are among the specific sub-population MSCs in the bone marrow, and they exhibit unique characteristics that help advance MSC-based cell therapy: (1) high proliferation potency, (2) multiple differentiation potencies (adipogenic, osteogenic, and chondrogenic differentiation), (3) low expression of senescence marker SA-beta-gal, and (4) uniform and small size, which allows them to avoid being trapped in the lung capillaries after intravenous administration in a mouse model [10]. Relatively rapidly proliferating clones were observed in single-cell sorted LT cells; they were named “rapidly expanding clones (RECs).” RECs are now expected to enhance the clinical trials for the treatment of hypophosphatasia and spinal canal stenosis [11], because their proliferative potency greatly assists in establishing a cell bank with less heterogeneity.

The expectations for the advancement of MSC-based therapeutic products have grown because MSCs are being used in leading translational studies for clinical applications [3, 12–14]. Thus, there is currently a great demand for the development of enabling technologies that can assist MSC manufacturing [15, 16]. The characteristics of RECs can greatly benefit from reducing two major critical risks in the present MSC manufacturing process to ensure efficient and robust cell manufacturing:

- (1) Difficulty in controlling quality variations among donors: MSCs have considerable donor variations [17–19], and it is practically impossible to examine sufficient variations in patient cells before developing a robust manufacturing process [20–22]. Thus, handling various unknown donor cells is a considerably risky task in MSC manufacturing, and it is the major cause of unexpected errors that can be dif-

ficult to solve. Allogenic cell bank establishment is a practical approach to control starting cell quality, which enables efficient MSC manufacturing. The REC can greatly benefit the process of establishing an allogenic cell bank. RECs can be mass produced by cell sorting, compared to difficulties involved in running large-scale donor selection until an ideal cell bank is obtained; therefore, the selection of the REC is more feasible and efficient. Furthermore, the proliferation potency of RECs considerably assists the success rate of establishing a cell bank while maintaining the banked cells in the earlier culture period. Thus, REC-based cell manufacturing enables more feasible process development and stable quality management.

- (2) Quality decay during cell expansion culture: MSCs lose their proliferation potency after several passages; in addition, other important quality attributes also degrade during expansion [23–31]. However, achieving a certain cell number is an essential quality criterion in the manufacturing of MSC-based therapeutic products. From a therapeutic perspective, the common protocol of MSC-based therapy requires more than a billion cells per treatment to ensure efficacy [32]. From a manufacturing aspect, the cells in the final cell bank must reach a large cell number to thoroughly test the final product with multiple quality criteria because most tests are invasive and the cells are consumed for each test [33]. However, such a cell expansion process is an unpromised trial with a high risk of failure owing to the quality decay probability. As a fact, the failure of the cell bank establishment can only be found “as a fact” after all the work, and it is difficult to avoid such failure beforehand. Within this context, the high and long-lasting proliferation potency of REC, which enables more than 10 repeats of serial passage, can greatly minimize the risk of cell bank establishment failure and realize the maximum culture efficiency for establishing a rich number of stocks in the earlier passages.

There is a dilemma in the establishment of cell banks for RECs despite the advantageous features of RECs that enable both the low-risk and high-efficiency manufacturing of allogenic therapeutic products. RECs can be produced effectively by cell sorting even from a small volume of cell source, and therefore, the starting clone variation can be greater than the conventional donor waiting for the allogenic cell banks. However, it is expensive to expand many RECs until the stage of the final cell bank; furthermore, final quality tests for assuring the established cell bank can incur further cost and effort. The variation in the starting RECs can trigger the expectation of selecting “better RECs” during the establishment of the cell bank; such expectations can increase the number of multiple cell bank establishments, which are cost- and effort-intensive. In practice, RECs are screened from LT cells using their primary growth speed from a single clone in a 96-well plate; all candidates are further expanded to form multiple “candidate cell banks” to select “the better REC cell bank” with higher quality. Such greedy selection for the better-quality product (= final product with lower risk and higher potency) is possible with RECs; however, it can raise the cost of the total process. Therefore, it is necessary to determine such candidate RECs as early as possible to help minimize excessive work. However, it is extremely difficult to test early-stage cells in the cell bank establishment, especially with single-clone-derived RECs, using conventional cell evaluation methods because most of them are invasive and waste precious cell sources.

In this study, we develop a morphology-based noninvasive potency prediction method for selecting “the better RECs” in the early stage of cell bank establishment where the cell number is extremely limited. Based on our previous findings on morphology-based early prediction for MSC quality decay [34–37], we attempt to predict the “further serial-passage potencies” using only their early morphological information to enhance the selection of the better RECs that form the better REC cell bank (Fig. 1). Furthermore, we expect that such potency prediction can help aid the cell bank establishment process to balance the “bank size” and “potency of banked cells” because it is also a critical dilemma for processing the cell bank. Manufacturing efficiency can be increased if a cell bank is largely expanded; however, this increases the risk of losing the proliferation potency of the banked cells. Therefore, we hypothesized that if future serial-passage potencies can be predicted in advance, it will enable to design a low-risk timing to achieve the maximum-sized cell bank which has high potency.

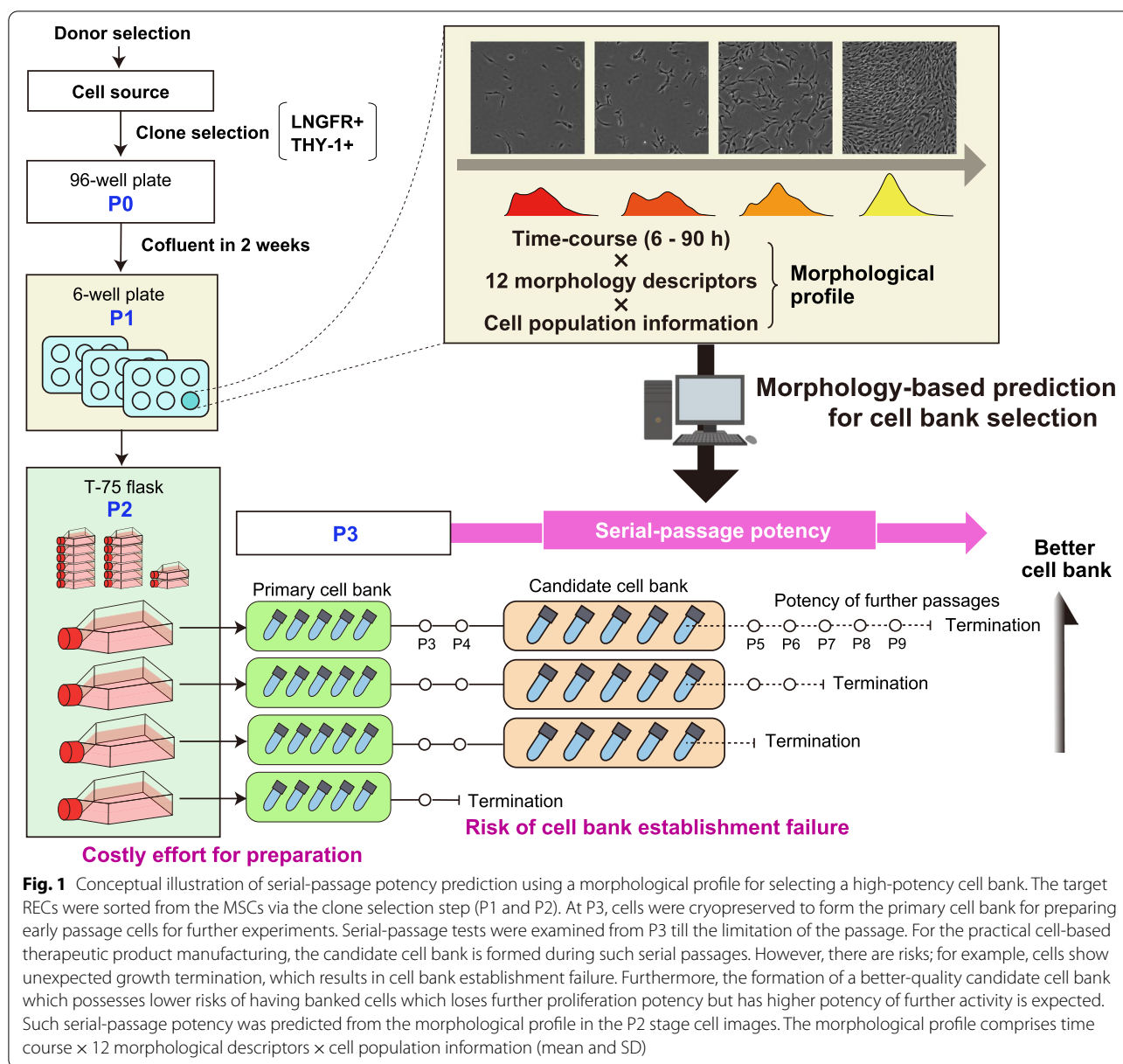
For the training data to develop a prediction model, we established 15 RECs from the bone marrow and experimentally confirmed the serial passage number until passage limitation (defined as “serial-passage

potencies”). Using the morphological descriptors in the stage of passage 1 (P1), we attempted to develop machine learning models to quantitatively predict such potency. During the development of prediction models, we conducted a detailed analysis of the data usage effect to obtain high-performance prediction models robustly. Thus, our data show the most effective time-course data usage and the minimum number of images required to realize the prediction model in a practical manner. The future potency prediction concept of our morphology-based REC indicates the potential of an image-based data-driven cell bank construction process in MSC manufacturing that can achieve both efficiency and robustness.

Methods

Cells and culture

Bone marrow mononuclear cells were prepared from bone marrow aspirate (AllCells, Alameda, USA) collected from healthy donors using density gradient centrifugation with Ficoll (GE Healthcare, Chicago, USA) to obtain RECs. Bone marrow mononuclear cells were stained with anti-human rabbit anti-CD90 IgG (BD Biosciences, Cat#559,869, Franklin Lakes, USA) and anti-human mouse anti-CD271 IgG (Thermo Fisher) for 1 h at 37 °C. Single-cell sorting was performed for CD90 and CD271 double-positive cells using the cell sorter (JSAN, BayBioscience, Kobe, Japan). The LT cells were sorted in 96-well plates (Thermo Fisher Scientific, Waltham, USA) as single clones (passage count = P0); they were further cultured with a maintenance medium (low glucose Dulbecco’s modified Eagle’s medium [DMEM] (Wako, Osaka, Japan) containing 20% fetal bovine serum [FBS] (Cytiva HyClone, Marlborough, USA), 20 ng/ml basic fibroblast growth factor [bFGF] (KAKEN PHARMACEUTICAL, Tokyo, Japan), 0.01 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid [HEPES] (Thermo Fisher Scientific, Waltham, USA), and 1% penicillin/streptomycin (Meiji Seika Pharma, Tokyo, Japan)). After 2 weeks, the clones were sub-confluent in 96-well plates, harvested, sub-cultured in a single well of a 6-well plate (Thermo Fisher Scientific) (passage count = P1), and set for image acquisition. Image acquisition for the further morphological analysis is done at this stage. The cells were then harvested and sub-cultured in T75 flasks (Thermo Fisher Scientific) (passage count: P2) when cells reached sub-confluency in a 6-well plate (Thermo Fisher Scientific). The cells were harvested using TrypLE Select (Thermo Fisher Scientific) after incubation for 3–5 min at 37 °C. The P2 cells in the T75 flasks that reached sub-confluent status were cryopreserved with 1.0×10^6 cells/ml with CP-1 (KYOKUTO PHARMACEUTICAL, Tokyo, Japan) and 25% human serum albumin (CSL Behring, Tokyo, Japan); they were designated “primary cell bank (P3).” With the primary cell



bank, in vitro and in vivo tests for the multiple potencies were excluded to save cells for the serial-passage experiments. For comparing the morphologies, conventionally processed bulk MSCs (BMMSCs: Lot 0,000,394,413, 0,000,411,107, 0,000,413,042, 0,000,422,610, 0,000,423,370, 0,000,429,365, 0,000,446,319, 0,000,451,491, 0,000,458,207) purchased from Lonza Japan, Ltd. (Tokyo, Japan) were cultured to P3 in MSCGM (Lonza Japan, Ltd.) supplemented with BulletKit (Lonza Japan, Ltd.). All cell cultures were maintained at 37 °C and 5% CO₂; the medium was changed once every 3 days.

Serial-passage potency test

The primary cell bank (P3) vial was thawed and seeded in a 100-mm dish (Thermo Fisher Scientific) with a density of 2.0×10^5 cells/dish, which we named as P3 data. Cells were harvested from a 100-mm dish using a TrypLE Select (Thermo Fisher Scientific) and sub-cultured in the new 100-mm dish with the same seeding density when the cells reached sub-confluent status. This sub-culture process was repeated until the cell growth was arrested. We counted the passage numbers and designated the final passage number as a “serial-passage limitation number” that we defined as the “serial-passage potency”

(Fig. 1). At each passage, the harvested total cell number was counted using Cellometer Auto T4 (Nexcelom Bioscience, Lawrence, USA), and the result was recorded as the total number of collected cells.

Image acquisition and processing

Phase-contrast microscopic images were acquired every 6 h for P1 RECs in a 6-well plate (Thermo Fisher Scientific) using the automatic image acquisition system Bio-Station CT (Nikon Corporation, Tokyo, Japan) at $\times 4$ magnification (64 tiling per well, covering 16 mm^2 , 1000 pixels/image). Each clone lot was imaged in one well; the cell number was extremely limited at this stage. For bulk MSCs, the MSCs were seeded into 6-well plates (Corning Incorporated, NY, USA) at 2000 cells/cm^2 ($n=3$ wells per lot). The time points were designated as time 1 (6 h after seeding) to 20 (120 h after seeding). Time 15 was selected as the final time point for morphology and growth rate measurements because some lots reached more than the sub-confluent status; therefore, it was difficult to recognize individual cells in the image accurately. Raw images were processed to measure individual cell morphologies for obtaining a summary of morphological profiles of the cell population (minimum of 300 cells to a maximum of 100,000 cells collected per well covered by 64 images). Image processing was performed by original codes using Python version 3.7.3, with packages NumPy version 1.20.0 and OpenCV version 4.4.0. The image processing pipeline was designed with eight processes: (1) background adjustment, (2) enhancement of texture, (3) binarization, (4) removal of small objects, (5) erosion, (6) removal of small objects, (7) fill hole, and (8) removal of frame-touching objects (Supplementary Fig. 1). After image processing, 12 basic morphological descriptors (Supplementary Table 1) were measured per cell region in each image. The summary of cell population was described by calculating “mean” and “standard deviation (SD)” using all single-cell-based morphological descriptor data. This summary represented nearly 1×10^4 to 1×10^5 cells from a single well covered by 64 images. We accumulated this morphological descriptor summary throughout the time course (6–90 h) and designated them as the “morphological profile” of the sample. The complete morphological profile for each condition (=1 well) comprised 360 parameters (=12 descriptors \times (mean and SD) \times 15 time points). All data processing was conducted using the R version 4.0.2.

Visualization of morphological profiles

Gaussian kernel density estimation was calculated using single-cell data at each time point to visualize the cell population using a single morphological descriptor. The distribution was estimated using a kernel function in R.

Raw data were transformed with log10 for distribution estimation to describe the “area.” Single-cell measurement data that exceed $\text{area} > 200$ pixels were collected to allow a detailed “area” discussion between RECs and conventionally processed bulk MSCs. Most objects smaller than 200 pixels were found in “round cells” during their proliferation, and therefore, it was considered difficult to discuss the size difference because the characteristics of expanding cells diminish in such data. The morphological profiles of all lots were analyzed using principal component analysis (PCA) to visualize the relative similarities of clones using multiple descriptors. Dots were colored with clone labels (15 colors) and serial-passage limitation numbers (gradations of seven levels: 7 to 13) in the PCA that compares clone morphological profiles. All data with the same time-window size, including different FOV number usages, were merged and used to set the total principal components for covering total data diversity to visualize the data usage effect and change their time-window size and numbers of FOVs. Then, data using different FOV numbers were plotted individually in the fixed PCA axis. In the comparative PCAs for the data-usage effect, one plot indicates “one clone.” The total single-cell measurement data from each data size (varying the combination of different time-window size and FOV numbers) were resampled by bootstrap (50 repeats allowing overlaps) to visualize the explanatory power of different data usages; their new mean and SD were re-obtained from each resampled data. A total of 50 plots per clone were plotted using PCA and such resampled data. Student’s *t*-test was used to test differences between the morphological descriptors and the population distributions of morphological descriptors. All data processing was performed using R (version 4.0.2).

Construction of prediction models for serial-passaging potency

Time-course morphological profiles were used as explanatory parameters, and an experimentally determined serial-passage limitation number was used as the objective parameter in the dataset for machine learning. Two types of machine learning models were examined—the linear regression model least absolute shrinkage and selection operator (LASSO) and the nonlinear machine learning model random forest (RF). With LASSO, parameter selection was performed using a mean decreasing Gini index. Model performances were validated by leave-one-out cross-validations and compared by root-mean-squared error (RMSE). The data usage effect in the morphological information was examined with an exhaustive combination of two parameters: length of time window (ranging from 6–90 h) and number of FOVs (ranging from 1 to 64). The total 6–90 h window was

shortened from the last time point (90 h) at each time point (6 h) to vary the length of the time window. Thus, the maximum total morphological descriptor comprises 360 parameters, whereas the minimum comprises 24 parameters. Images from 64 images were selected randomly to vary the number of FOVs. The bootstrapping of 50 repeated times of re-sampling FOVs from 64 images was introduced to increase the dataset size from 15 to 750 samples for evaluating the data variation effect on selected data size conditions (6 h, 6–30 h, 6–60 h vs. 15, 40, and 60 FOVs). All data processing and machine learning were performed using the R version 4.0.2.

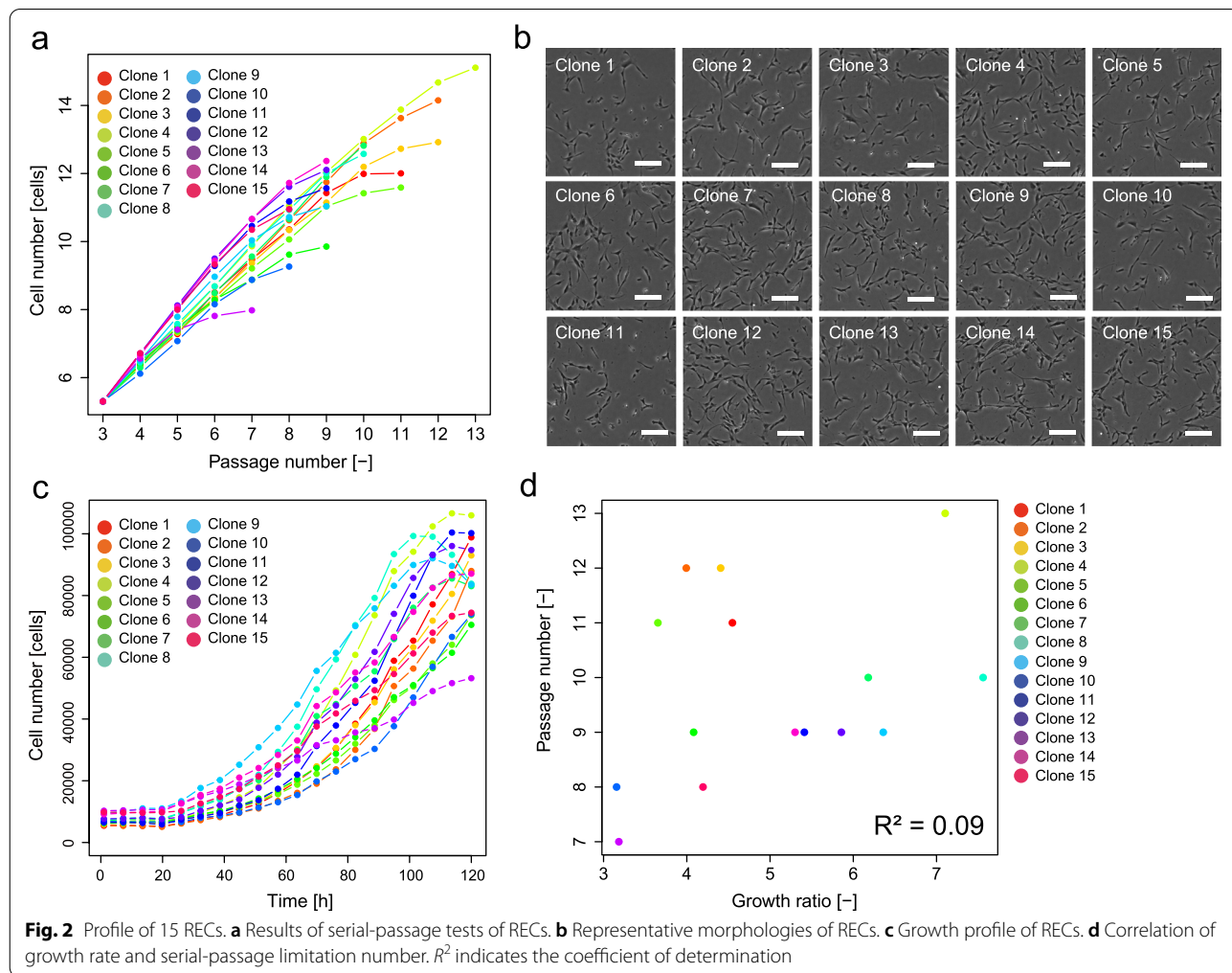
Results

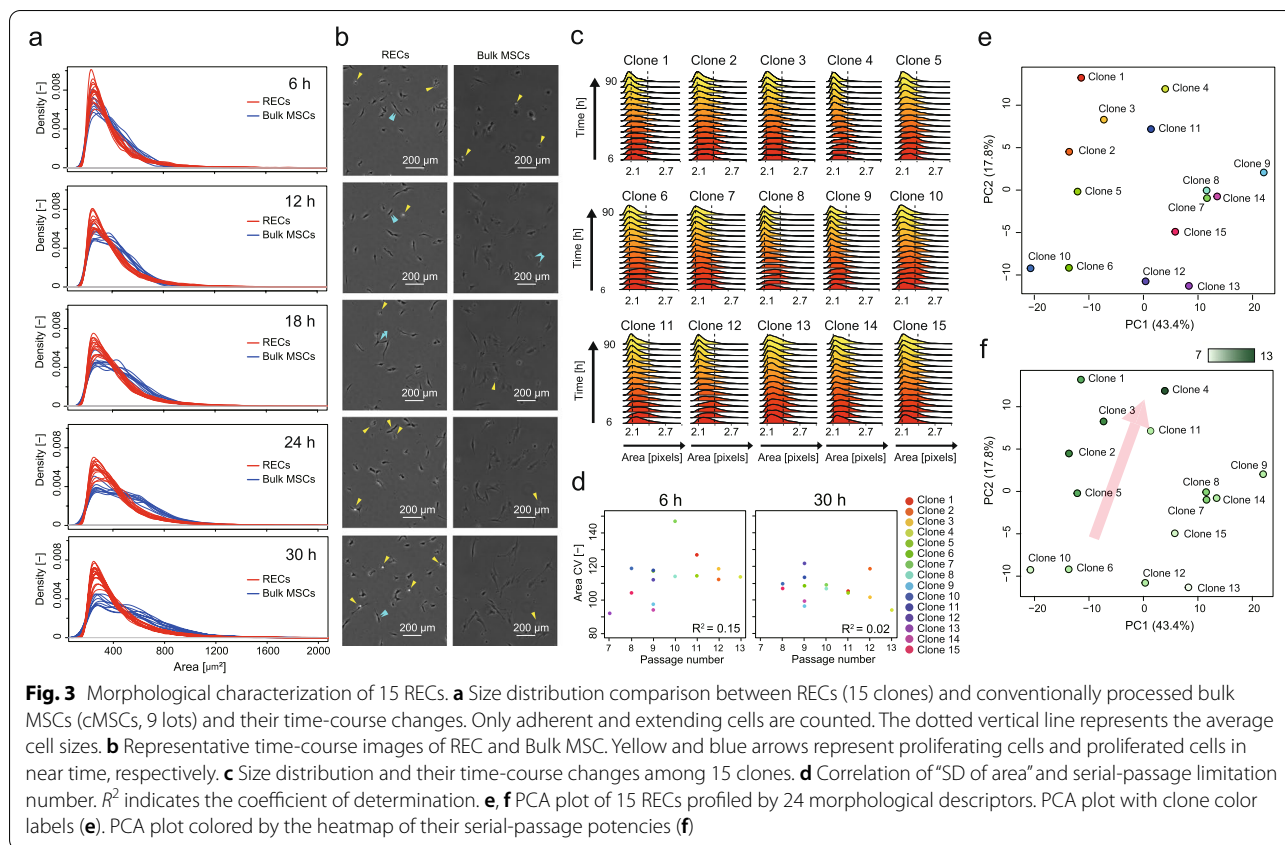
Collection and characterization of RECs for the training data

We started our work to achieve 15 clones of RECs from a single donor of bone marrow mononuclear cells (BMMCs) to develop a machine learning model for predicting “serial-passage potency” from the early-stage cell

morphologies (Scheme in Fig. 1). Fifteen clones were collected from the same donor sample and sorted according to the basic criteria for REC: (1) LNGFR and THY-1 double-positive in cell sorting and (2) reaching sub-confluent status within 2 weeks after single-clone sorting in 96-well plates (P0 stage). Then, the clones expanded in a 6-well plate followed by a T-75 flask were cryopreserved as primary cell banks (P3), which is the bank for storing early passage cells.

We evaluated the serial-passage potencies of candidate clones in P2 to further select “the better RECs.” The cellular performance that enables several repeated rounds of passages can be considered an ideal criterion for selecting “a better-quality cell bank” for further usage. The serial-passage potency test results indicated that our RECs retained high proliferation potencies, which marks 9.8 repeats of serial passage on average (Fig. 2a). Such a potency can be considered high compared to that of commercially available MSCs processed using conventional methods. However, some





clones lost their proliferation gradually, even among clones that passed the REC criteria and retained long serial-passage potencies (e.g., clone 11 or 13 in our data).

Next, we evaluated the morphological characteristics of the clones (Fig. 2b). At the P1 stage in 6-well plates, which is an extremely early stage, the morphology includes an important signature of the cells for their evaluation. However, manual morphological observation (without quantification) makes it difficult to discriminate the differences between the RECs.

We analyzed the growth profiles of RECs at the P1 stage using time-course images (Fig. 2c). Among the 15 clones, 12 showed a growth rate over fourfold; this growth rate was significantly higher than that of several conventionally processed bulk MSCs in our study (Supplementary Fig. 2). The clone with the lowest serial-passage potency (clone 13) showed a low growth rate; the highest serial-passage potency clone (clone 4) showed a high growth rate. However, the coefficient of determination between the “growth rate” and the “serial-passage limitation number” was low ($R^2 = 0.09$) (Fig. 2d). Thus, the data indicate that the growth rate measurement at the P1 stage cannot predict future serial-passage potencies.

Morphological characterization of RECs

We attempted to characterize RECs (at P1) quantitatively via image-based morphological analysis according to our previously reported analysis concepts [33–36]. Compared with conventionally processed bulk MSCs (cMSCs, nine lots), RECs were more homogeneous and remained smaller even after adhesion (Fig. 3a, b). Both RECs and cMSCs started from a similarly sized population (median area = $358 \mu\text{m}^2$, $335 \mu\text{m}^2$, respectively) at the very early adhesion stage (6 h, T -test $p < 0.22$); the median size of the RECs remained small ($363 \mu\text{m}^2$) wherein the bulk cMSCs expanded during the culture after 30 h (median area = $473 \mu\text{m}^2$, T -test $p < 0.000001$). Furthermore, more proliferating cells that are visually white and round under phase-contrast microscopy were found in the RECs during the same series of time-course images (Fig. 3b).

We found that the adherent cell population remained broad in clones with decreased serial-passage potencies by visualizing the population change of such RECs during the time course at P1 stage (Fig. 3c). Thus, high-potency cells can achieve a homogenous size population; however, such size data analysis was suggestive, and the coefficient of determination between the “SD of area” and the “serial-passage limitation number” remained low

(Fig. 3d). A single morphology descriptor analysis was not sufficiently efficient for a quantitative prediction.

We profiled clones using multiple morphological information described with 24 descriptors obtained from the basic descriptors (Supplementary Table 1) [33–36]. The morphological similarities of the RECs are visualized using principal component analysis (PCA) (Fig. 3e is labeled by clone numbers, and Fig. 3f is labeled by the serial-passage-potency of each clone). These results indicate that there are certain clusters of clones with similar morphological profiles that slightly divide the low- and high-serial-passage potency clones. In practice, low serial-passage potency clones gather in the low PC2 axis, whereas the high serial-passage potency clones gather in the high PC2 axis and in the middle of the PC1 axis. The contributing descriptors in the axes of the PCA map, especially in the most explanatory PC2 axis, can be interpreted as follows (Supplementary Table 2): The clone has a longer serial-passage potency when cells are more homogeneous and show a spindle shape during the 24–30-h growth; however, their serial-passage potency is shorter when cellular morphological homogeneity is disturbed. Such an unsupervised model analysis suggests that a multiple-descriptor combination provides a better explanation to morphologically characterize potencies.

Morphology-based machine learning for predicting serial-passage potency

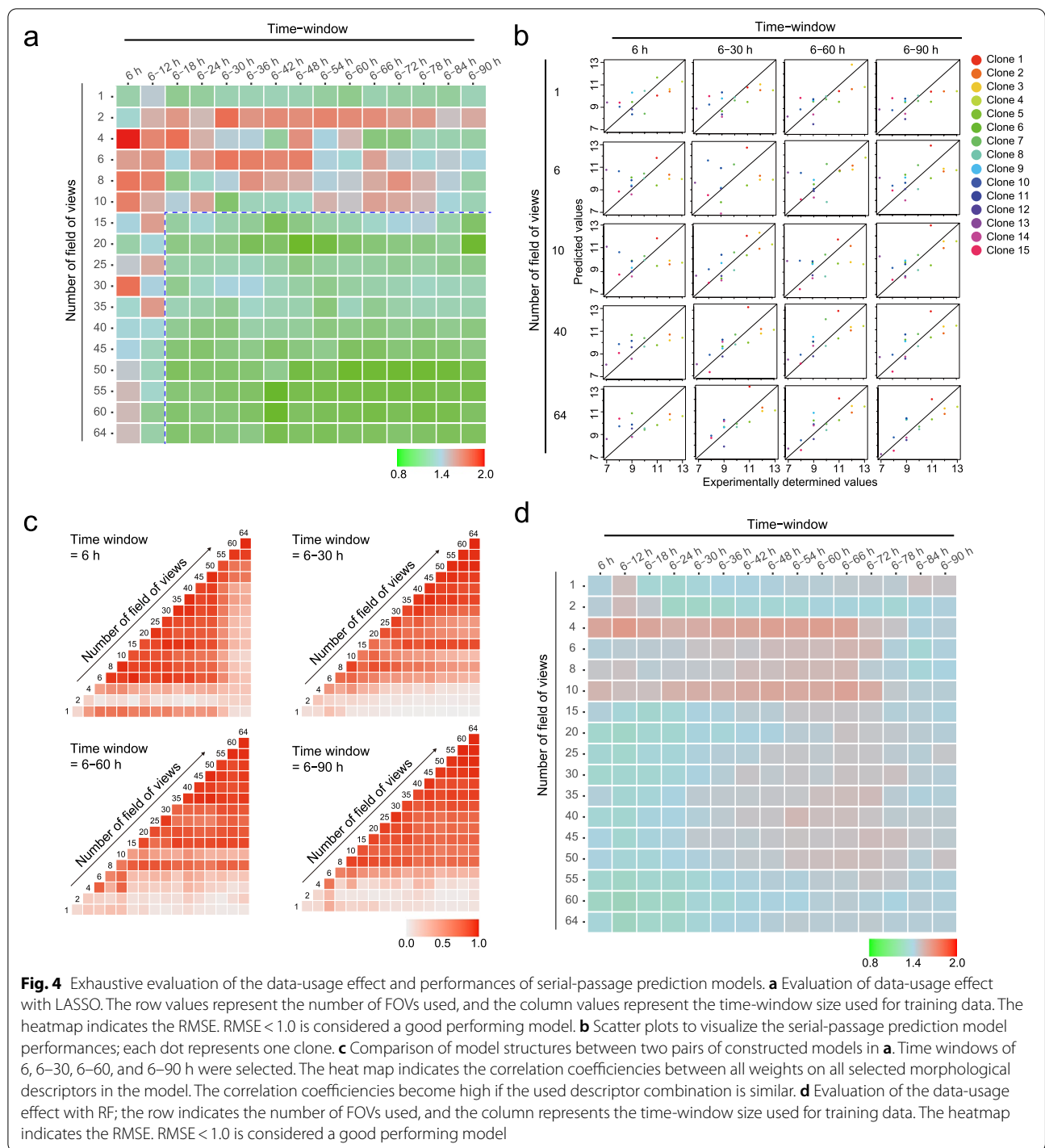
We next investigated the development of machine learning models with morphological information to enable the quantitative prediction of “serial-passage potencies” of RECs based only on morphological information (Fig. 1). We used morphological profiles (24 descriptors \times 15 time points) to predict the serial-passage limitation numbers. Under this prediction model development, we attempted to understand the part or the extent to which the morphological information effectively contributes to the development of the prediction model. Thus, we investigated the effect of morphological data usage by changing two parameters: the time-window effect and the field-of-view (FOV)-number effect. We investigated these parameters because we found that the image data collection effort can be effectively reduced via such detailed investigation in our previous challenge involving the prediction of the growth rate and osteogenic differentiation rate of MSCs from morphological descriptors [33–36]. Shortening the time window and minimizing the FOV number can not only save time and effort for image data acquisition, but can also accelerate the prediction.

From the exhaustive examination of both the “time-window effect” and the “FOV-number effect” with least absolute shrinkage and selection operator (LASSO), we found that high-performance prediction models

(RMSE $<$ 1.0 shown as green colored heatmap in Fig. 4a) can be obtained with several parameter combinations even with the morphological information in P1 stage cells. The data suggest that the prediction performance can be maintained even if the time window and FOV numbers are reduced. Furthermore, it was deduced that the effect of “FOV-numbers” is more important than that of the “time window” because the time window can be shortened without performance degradation when more than 15 FOVs are collected.

A scatter plot is plotted to further understand the model performance for predicting serial-passage potencies for each clone (Fig. 4b). These data clearly show that our prediction models predict “quantitative values of serial-passage limitation numbers.” Many high-performance prediction models have been developed using different sizes of training data (Fig. 4a, b). However, we suspected that all model structures were randomly different by lacking a common structure because LASSO is the algorithm that creates the best descriptor combination for each dataset; similar performance models with a completely different model structure can be obtained. If the model structure is different in each condition of the data, such modeling result is not robust, and therefore not practical. Thus, we compared all model structure correlations (Fig. 4c) and confirmed that the top-performing models robustly share similar model structures. This result suggested that the relation between the morphological descriptor combination and serial-passage potencies can be modeled with a certain universal combination of morphological descriptors.

A comparison between the highly contributing descriptors shared between different models, i.e., “Correlation_SD (18 h),” “Correlation_SD (6 h),” and “Energy_SD (18 h),” contributed positively to predict the long serial-passage potency clones; “Correlation_mean (6 h),” “Length_SD (18 h),” and “Compactness_SD (18 h)” contributed negatively to predict the short serial-passage potency clones (Supplementary Table 3). Both “correlation” and “energy” are texture descriptors, and therefore, they commonly reflect the three-dimensional pattern and complexity of cells because it changes the intensity profile under phase-contrast microscopy. In practice, the intensity profiles change drastically when cells change their three-dimensional roundness during cell division. Thus, an increase in the “SD of texture” indicates that there are more proliferating cell populations. Thus, the population becomes more homogeneous in texture when the “mean of texture” has a greater effect; this implies that there is a decrease in the number of cellular events that change texture. Length and compactness are shape-related descriptors; therefore, they commonly reflect



two-dimensional responses such as elongation and expansion. If such a shape changes with few textural changes, it indicates that the cells use their activity for elongation more instead of for proliferation. Thus, such model structure-derived information suggests that the acquired serial-passage prediction models are not

only useful for early detection of their future potency, but are also informative for the quantitative extraction of the morphological rule, which is descriptive and recordable. Thus, such a descriptive understanding of morphological profiles helps escape from the old habits of grasping morphological changes by feeling.

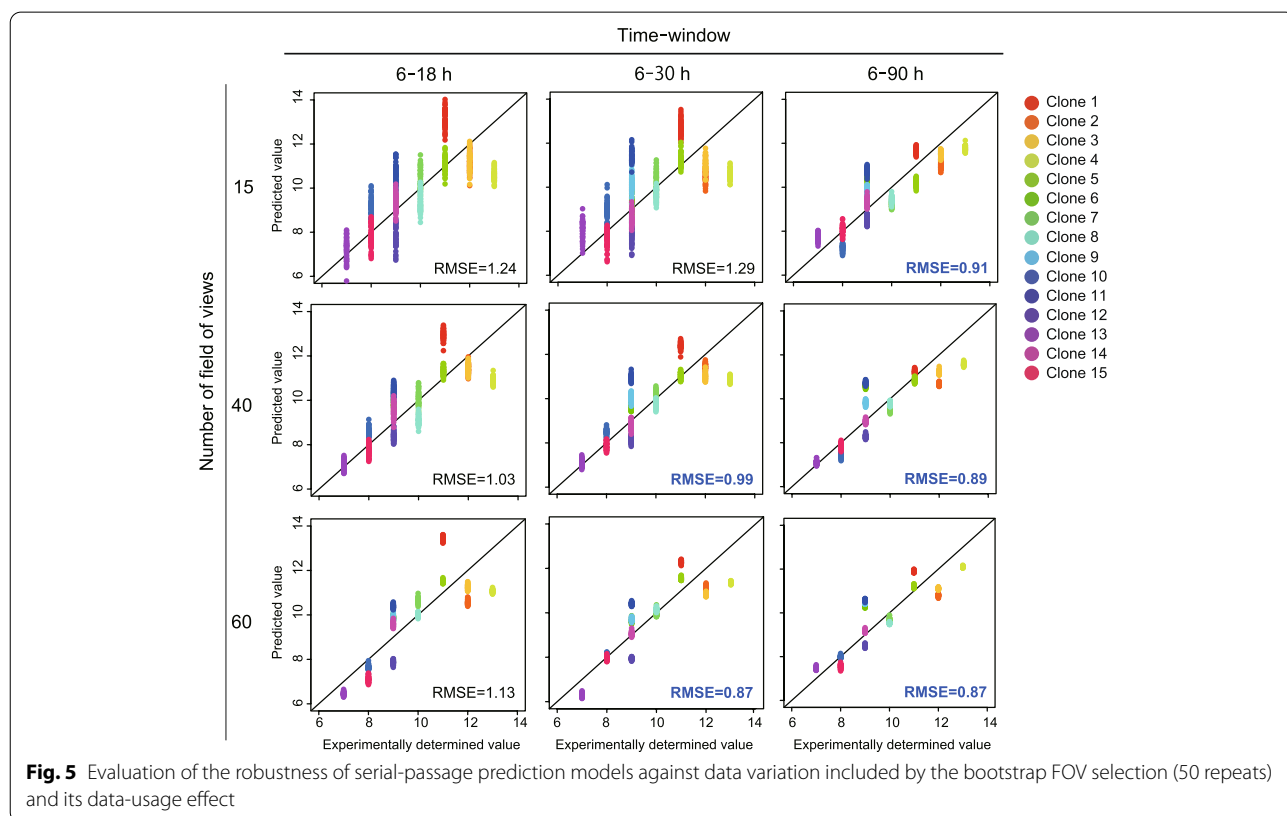
The model did not show a better performance than LASSO in any data usage conditions when we examined the same condition matrix with the nonlinear machine learning model RF (Fig. 4d). These data reflect that the serial-passage potency and morphological information are linearly related.

Finally, we attempted to confirm the predictive performance of LASSO models in more detail because the model training was based on their raw sample data, which are relatively small in size compared to other fields' machine learning applications, although we examined two parameter combinations (time window and FOV number). We introduced bootstrap to increase the variations of image-derived morphological profiles for evaluating the extent to which the prediction model can stand robustly to the effect from the data variation. We examine the performance of the prediction models by introducing 50 bootstraps to collect different FOV combinations from the 64-tiling images per sample (Fig. 5). The result indicates that such image-sampling bias introduced by bootstrapping caused performance degradation in some prediction models. However, although the range of the "time-window effect" and "FOV-number effect" for achieving high-performance models (RMSE < 1.0) was narrowed by a more robust model compared to that in Fig. 4a–c, we can minimize the data collection size

to 6–30 h with 40 FOVs for the earliest prediction, and 6–90 h with 15 FOVs for the minimum effort prediction while keeping the prediction accuracy (RMSE < 1.0). The FOV number clearly improved the morphological profile robustness when we evaluated such a bootstrap effect on PCA; the accumulation of the morphological profile by a longer time window improved to discriminate between clone differences. Such data-size effect investigations will contribute to designing a more effective process of introducing an image-based quality check in cell bank establishment.

Discussions

RECs are clonal MSCs selected from human BMMCs, which not only retain the superior qualities of conventionally processed MSCs, but also have characteristics that are advantageous for the practical cell manufacturing process for therapeutic products. In particular, the high proliferative potency of RECs provides a major advantage in developing efficient manufacturing processes for cell therapeutic products. Therefore, in this study, we investigated the possibility of predicting continuous passaging capacity from initial morphological information alone and its most practical construction method so that the capacity of RECs can be evaluated from the initial stage of cell bank construction.



A long unsolved problem in any type of cell culture is determining “the best timing to make cryo-stocks” in the expansion culture. Since most normal cells lose their proliferation potency when cultured *in vitro* [38, 39], one can only bet on which passage number to end with while making cryostocks. For establishing industrial cell banks, such a betting factor amounts to a significant risk: the low-success-rate expansion culture will incur significant expenses if the cell does not proliferate as expected, and if the cells are cryopreserved too early, the stock will not profit production efficacy. In practice, the bottleneck in the practical cell bank establishment is the effort of recruiting precious donors, and not the effort of making greedy selection of candidate cell banks. However, with RECs, we expect a stricter and more selective process for finalizing the candidate cell bank as a “master cell bank.” Our investigation presents a new concept of using morphological noninvasive analysis as an “in-process analysis tool” for enhancing and optimizing the cell bank establishment process. This concept will help set the best cryostock production timing by balancing “the yield of cells” and “the remaining potency of banked cells” and predicting the future serial-passage potency. Such an approach will help discard the present cell bank design concept, which restricts the passage number using data-less logic.

Although we investigated a method to predict the “serial-passage potency” of RECs, but the continuous passage potency in cMSCs may require some discussion. It is understood that for the induced pluripotent stem cells (iPSCs), the uncontrollable proliferation potency in iPSCs has a negative effect on clinical treatment, for example, the risk of teratoma formation [40, 41]. However, with MSCs, which are known to exhibit limited proliferation potency, their serial-passage potencies are considered with several aspects. If the “candidate cell bank” of this study is prepared as a master cell bank for the creation of further working cell banks, its serial-passage potency would be beneficial to the entire process. However, the serial-passage potency effect should be carefully examined if it is prepared as the final cell bank for implantation. If it profits the efficacy of the final product, it can be an advantage; however, if it negatively affects the efficacy, it will be a risk. In any case, such future potency prediction from the earliest stage of the process will help optimize the final cell bank quality because it can only be evaluated by excessive continuous evaluations for futile REC candidates. Since REC is currently on the path to clinical trials, our next challenge is to validate the effectiveness of such potency predictions and efficiently move forward with product manufacturing.

Finally, our morphology-based future potency prediction on RECs triggers the discussion of whether this developed model can be applied to other MSC cell-bank

establishment studies. Currently, we consider that our model is still limited to predicting and evaluating RECs. This interpretation is not attributed to the difference in potencies in RECs and bulk MSCs because we clearly found that the morphological distribution of RECs differs from that of bulk MSCs. Our image-based detailed morphology measurement indicated that the major population of RECs comprises nearly twofold smaller cells compared to the bulk MSCs. Such large size differences would make the prediction model structure, the combination of morphological descriptors, fit for RECs because we use “mean and SD” for reflecting the cell population distribution for our morphological profiles. Furthermore, it is practically difficult for bulk MSCs to conduct “serial-passage tests” for as long as RECs. Such large differences between the native potencies of RECs and bulk MSCs can result in unexpected data bias that can unexpectedly develop serial-passage prediction models that discriminate “RECs or bulk MSCs” from biased data. Therefore, it is a future challenge to investigate such universal morphological characteristics that can be reflected in other stem cells.

Conclusions

Although our findings are based on a limited number of clones, our investigation of image-based machine learning models was found to introduce a new concept of data-driven process management for a more effective cell bank establishment. Our next challenge will be to expand our morphology-based early cell potency predictions to obtain clones with higher differentiation potencies which closely relate to the therapeutic effects.

Abbreviation

MSC: Mesenchymal stem cell; BMMC: Bone marrow mononuclear cell; REC: Rapidly expanding MSC clone; P1: Passage number 1; P2: Passage number 2; P3: Passage number 3; SD: Standard deviation; PCA: Principal component analysis; PC1: Principal component 1; PC2: Principal component 2; LASSO: Least absolute shrinkage and selection operator; RF: Random Forest; FOV: Field-of-view; RMSE: Root-mean-squared error; iPSC: Induced pluripotent stem cell.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41232-022-00214-w>.

Additional file 1: Supplementary Fig. 1. Schematic pipeline of image processing used in this study.

Additional file 2: Supplementary Table 1. Basic morphological parameters measured per cell.

Additional file 3: Supplementary Fig. 2. Growth profiles of conventionally processed bulk MSCs (nine lots).

Additional file 4: Supplementary Table 2. Morphological descriptors highly contributing to PCA.

Additional file 5: Supplementary Table 3. Morphological descriptors used in LASSO models.

Acknowledgements

We thank Kenji Ito (Inotech Corp.) for assisting the cloud linkage between Shimane University and Nagoya University for enabling image data sharing with the developed cloud system in the NEDO project.

Authors' contributions

Conceptualization, T. S., Y. T., R. K., and Y. M.; investigation, T. S., Y. T., H. M., Y. K., and R. K.; writing — original draft, T. S., Y. T., and R. K.; writing — review and editing, R. K., and Y. M. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Next Generation AI Technology Field: Strategic Advancement of Multi-Purpose Ultra-Human and AI Technologies (Samurai) from the New Energy and Industrial Technology Development Organization (NEDO) to RK and YM.

Availability of data and materials

Datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

TS, HM, and YK are employees of PuREC, whose main business product is REC cells. YM is the CTO of PuREC, whose main business product is REC cells, and owns a stock share.

Author details

¹Department of Life Science, Faculty of Medicine, Shimane University, 89-1 Enya-cho, Izumo, Shimane 693-8501, Japan. ²PuREC Co. Ltd, 89-1 Enya-cho, Izumo, Shimane 693-8501, Japan. ³Department of Basic Medicinal Sciences, Graduate School of Pharmaceutical Sciences, Nagoya University, Tokai National Higher Education and Research System, Furocho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan. ⁴Institute of Nano-Life-Systems, Institutes of Innovation for Future Society, Nagoya University, Tokai National Higher Education and Research System, Furocho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan.

Received: 7 April 2022 Accepted: 29 May 2022

Published online: 02 October 2022

References

- Marofi F, Vahedi G, Biglari A, Esmailzadeh A, Athari SS. Mesenchymal stromal/stem cells: a new era in the cell-based targeted gene therapy of cancer. *Front Immunol*. 2017;8:1770. <https://doi.org/10.3389/fimmu.2017.01770>.
- Sherman LS, Shaker M, Mariotti V, Rameshwar P. Mesenchymal stromal/stem cells in drug therapy: new perspective. *Cytotherapy*. 2017;19(1):19–27. <https://doi.org/10.1016/j.jcyt.2016.09.007>.
- Squillaro T, Peluso G, Galderisi U. Clinical trials with mesenchymal stem cells: an update. *Cell Transplant*. 2016;25(5):829–48. <https://doi.org/10.3727/096368915X689622>.
- Wakao S, Kitada M, Kuroda Y, Shigemoto T, Matsuse D, Akashi H, Tanimura Y, Tsuchiyama K, Kikuchi T, Goda M, Nakahata T, Fujiyoshi Y, Dezawa M. Multilineage-differentiating stress-enduring (Muse) cells are a primary source of induced pluripotent stem cells in human fibroblasts. *Proc Natl Acad Sci USA*. 2011;108(24):9875–80. <https://doi.org/10.1073/pnas.1100816108>.
- Vogel W, Grünebach F, Messam CA, Kanz L, Brugger W, Bühring HJ. Heterogeneity among human bone marrow-derived mesenchymal stem cells and neural progenitor cells. *Haematologica*. 2003;88(2):126–33.
- Trivanović D, Jauković A, Popović B, Krstić J, Mojsilović S, Okić-Djordjević I, Kukulj T, Obradović H, Santibanez JF, Bugarski D. Mesenchymal stem cells of different origin: comparative evaluation of proliferative capacity, telomere length and pluripotency marker expression. *Life Sci*. 2015;141:61–73. <https://doi.org/10.1016/j.lfs.2015.09.019>.
- Sacchetti B, Funari A, Remoli C, et al. No identical “mesenchymal stem cells” at different times and sites: human committed progenitors of distinct origin and differentiation potential are incorporated as adventitial cells in microvessels. *Stem Cell Reports*. 2016;6(6):897–913. <https://doi.org/10.1016/j.stemcr.2016.05.011>.
- Bakopoulou A, Apatzidou D, Aggelidou E, Gousopoulou E, Leyhausen G, Volk J, Kritis A, Koidis P, Geurtsen W. Isolation and prolonged expansion of oral mesenchymal stem cells under clinical-grade, GMP-compliant conditions differentially affects “stemness” properties. *Stem Cell Res Ther*. 2017;8(1):247. <https://doi.org/10.1186/s13287-017-0705-0>.
- Colter DC, Class R, DiGirolamo CM, Prockop DJ. Rapid expansion of recycling stem cells in cultures of plastic-adherent cells from human bone marrow. *Proc Natl Acad Sci USA*. 2000;97(7):3213–8. <https://doi.org/10.1073/pnas.97.7.3213>.
- Mabuchi Y, Morikawa S, Harada S, Niibe K, Suzuki S, Renault-Mihara F, Houlihan DD, Akazawa C, Okano H, Matsuzaki Y. LNGFR+THY-1+VCAM-1hi+ cells reveal functionally distinct subpopulations in mesenchymal stem cells. *Stem Cell Reports*. 2013;1(2):152–65. <https://doi.org/10.1016/j.stemcr.2013.06.001>.
- Ukeba D, Yamada K, Tsujimoto T, Ura K, Nonoyama T, Iwasaki N, Sudo H. Bone marrow aspirate concentrate combined with in situ forming biore-sorbable gel enhances intervertebral disc regeneration in rabbits. *J Bone Joint Surg*. 2021;103(8): e31. <https://doi.org/10.2106/JBJS.20.00606>.
- De Bari C, Roelofs AJ. Stem cell-based therapeutic strategies for cartilage defects and osteoarthritis. *Curr Opin Pharmacol*. 2018;40:74–80. <https://doi.org/10.1016/j.coph.2018.03.009>.
- Kabat M, Bobkov I, Kumar S, Grumet M. Trends in mesenchymal stem cell clinical trials 2004–2018: is efficacy optimal in a narrow dose range? *Stem Cells Transl Med*. 2020;9(1):17–27. <https://doi.org/10.1002/sctm.19-0202>.
- Viswanathan S, Keating A, Deans R, Hematti P, Prockop DJ, Stroncek DF, Stacey G, Weiss DJ, Mason C, Rao MS. Soliciting strategies for developing cell-based reference materials to advance mesenchymal stromal cell research and clinical translation. *Stem Cells and Development*. 2014;23(11):1157–67. <https://doi.org/10.1089/scd.2013.0591>.
- Dwarshuis NJ, Parratt K, Santiago-Miranda A, Roy K. Cells as advanced therapeutics: state-of-the-art, challenges, and opportunities in large scale biomanufacturing of high-quality cells for adoptive immunotherapies. *Adv Drug Deliv Rev*. 2017;114:222–39. <https://doi.org/10.1016/j.addr.2017.06.005>.
- Jossen V, van den Bos C, Eibl R, Eibl D. Manufacturing human mesenchymal stem cells at clinical scale: process and regulatory challenges. *Appl Microbiol Biotechnol*. 2018;102:3981–94. <https://doi.org/10.1007/s00253-018-8912-x>.
- Lo Surdo J, Bauer SR. Quantitative approaches to detect donor and passage differences in adipogenic potential and clonogenicity in human bone marrow-derived mesenchymal stem cells. *Tissue Engineering - Part C: Methods*. 2012;18(11):877–89. <https://doi.org/10.1089/ten.tec.2011.0736>.
- Mohamed-Ahmed S, Fristad I, Lie SA, Suliman S, Mustafa K, Vindenes H, Idris SB. Adipose-derived and bone marrow mesenchymal stem cells: a donor-matched comparison. *Stem Cell Res Ther*. 2018;9(1):168. <https://doi.org/10.1186/s13287-018-0914-1>.
- Fennema EM, Renard AJS, Leusink A, Van Blitterswijk CA, De Boer J. The effect of bone marrow aspiration strategy on the yield and quality of human mesenchymal stem cells. *Acta Orthop*. 2009;80(5):618–21. <https://doi.org/10.3109/17453670903278241>.
- Collart-Dutilleul P-Y, Chaubron F, De Vos J, Cuisinier FJ. Allogenic banking of dental pulp stem cells for innovative therapeutics. *World Journal of Stem Cells*. 2015;7(7):1010–21. <https://doi.org/10.4252/wjsc.v7.i7.1010>.

21. Nievale S. Concise review: umbilical cord derived mesenchymal stem cell bank. *Progress in Stem Cell*. 2017;4:3–4. <https://doi.org/10.15419/psc.v4i3-4.397>.
22. Alzahrani FA, Saadeldin IM, Ahmad A, Kumar D, Azhar EI, Siddiqui AJ, Kurdi B, Sajini A, Alrefaei AF, Jahan S. The potential use of mesenchymal stem cells and their derived exosomes as immunomodulatory agents for COVID-19 patients. *Stem Cells International*. 2020;2020:8835986. <https://doi.org/10.1155/2020/8835986>.
23. Rombouts WJC, Ploemacher RE. Primary murine MSC show highly efficient homing to the bone marrow but lose homing ability following culture. *Leukemia*. 2003;17(1):160–70. <https://doi.org/10.1038/sj.leu.2402763>.
24. Dhanasekaran M, Indumathi S, Poojitha R, Kanmani A, Rajkumar JS, Sudarsanam D. Plasticity and banking potential of cultured adipose tissue derived mesenchymal stem cells. *Cell Tissue Banking*. 2013;14(2):303–15. <https://doi.org/10.1007/s10561-012-9311-7>.
25. Dave C, McRae A, Doxtator E, Mei SHJ, Sullivan K, Wolfe D, Champagne J, McIntyre L. Comparison of freshly cultured versus freshly thawed (cryopreserved) mesenchymal stem cells in preclinical in vivo models of inflammation: a protocol for a preclinical systematic review and meta-analysis. *Syst Rev*. 2020;9(1):188. <https://doi.org/10.1186/s13643-020-01437-z>.
26. Hladik D, Höfig I, Oestreicher U, Beckers J, Matjanovski M, Bao X, Scherthan H, Atkinson MJ, Rosemann M. Long-term culture of mesenchymal stem cells impairs ATM-dependent recognition of DNA breaks and increases genetic instability. *Stem Cell Res Ther*. 2019;10(1):218. <https://doi.org/10.1186/s13287-019-1334-6>.
27. Jin HJ, Bae YK, Kim M, Kwon SJ, Jeon HB, Choi SJ, Kim SW, Yang YS, Oh W, Chang JW. Comparative analysis of human mesenchymal stem cells from bone marrow, adipose tissue, and umbilical cord blood as sources of cell therapy. *Int J Mol Sci*. 2013;14(9):218. <https://doi.org/10.3390/ijms140917986>.
28. Yang YHK, Ogando CR, Wang See C, Chang TY, Barabino GA. Changes in phenotype and differentiation potential of human mesenchymal stem cells aging in vitro. *Stem Cell Res Ther*. 2018;9(1):131. <https://doi.org/10.1186/s13287-018-0876-3>.
29. Wagner W, Horn P, Castoldi M, Diehlmann A, Bork S, Saffrich R, Benes V, Blake J, Pfister S, Eckstein V, Ho AD. Replicative senescence of mesenchymal stem cells: a continuous and organized process. *PLoS ONE*. 2008;3(5):e2213. <https://doi.org/10.1371/journal.pone.0002213>.
30. Turinetti V, Vitale E, Giachino C. Senescence in human mesenchymal stem cells: functional changes and implications in stem cell-based therapy. *Int J Mol Sci*. 2016;17(7):1164. <https://doi.org/10.3390/ijms17071164>.
31. Binato R, de Souza FT, Lazzarotto-Silva C, Du Rocher B, Mencialha A, Pizzatti L, Bouzas LF, Abdelhay E. Stability of human mesenchymal stem cells during in vitro culture: considerations for cell therapy. *Cell Prolif*. 2013;46(1):10–22. <https://doi.org/10.1111/cpr.12002>.
32. Galipeau J, Sensébé L. Mesenchymal stromal cells: clinical challenges and therapeutic opportunities. *Cell Stem Cell*. 2018;22(6):824–33. <https://doi.org/10.1016/j.stem.2018.05.004>.
33. Samsanraj RM, Raghunath M, Nurcombe V, Hui JH, van Wijnen AJ, Cool SM. Concise review: multifaceted characterization of human mesenchymal stem cells for use in regenerative medicine. *Stem Cells Transl Med*. 2017;6(12):2173–85. <https://doi.org/10.1002/sctm.17-0129>.
34. Imai Y, Yoshida K, Matsumoto M, Okada M, Kanie K, Shimizu K, Honda H, Kato R. In-process evaluation of culture errors using morphology-based image analysis. *Regenerative Therapy*. 2018;9(9):15–23. <https://doi.org/10.1016/j.reth.2018.06.001>.
35. Matsuoka F, Takeuchi I, Agata H, Kagami H, Shiono H, Kiyota Y, Honda H, Kato R. Morphology-based prediction of osteogenic differentiation potential of human mesenchymal stem cells. *PLoS ONE*. 2013;8(2):e55082. <https://doi.org/10.1371/journal.pone.0055082>.
36. Sasaki H, Takeuchi I, Okada M, Sawada R, Kanie K, Kiyota Y, Honda H, Kato R. Label-free morphology-based prediction of multiple differentiation potentials of human mesenchymal stem cells for early evaluation of intact cells. *PLoS ONE*. 2014;9(4):e93952. <https://doi.org/10.1371/journal.pone.0093952>.
37. Takemoto Y, Imai Y, Kanie K, Kato R. Predicting quality decay in continuously passaged mesenchymal stem cells by detecting morphological anomalies. *J Biosci Bioeng*. 2020;131(2):198–206. <https://doi.org/10.1016/j.jbiosc.2020.09.022>.
38. Sawada R, Ito T, Tsuchiya T. 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*. 2006;9(3):179–84. <https://doi.org/10.1007/s10047-006-0338-z>.
39. Muraglia A, Cancedda R, Quarto R. Clonal mesenchymal progenitors from human bone marrow differentiate in vitro according to a hierarchical model. *J Cell Sci*. 2000;113(7):1161–6. <https://doi.org/10.1242/jcs.113.7.1161>.
40. Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature*. 2007;448(7151):313–7. <https://doi.org/10.1038/nature05934>.
41. Gutierrez-Aranda I, Ramos-Mejia V, Bueno C, Munoz-Lopez M, Real PJ, Mácia A, Sanchez L, Ligerio G, Garcia-Perez JL, Menendez P. Human induced pluripotent stem cells develop teratoma more efficiently and faster than human embryonic stem cells regardless the site of injection. *Stem Cells*. 2010;28(9):1568–70. <https://doi.org/10.1002/stem.471>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

