

Morphology-based analysis of myoblasts for prediction of myotube formation

Kiyoshi Ishikawa^{1,2†}, Kei Yoshida^{1†}, Kei Kanie¹, Kenji Omori², Ryuji Kato¹

¹Graduate School of Pharmaceutical Sciences, Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-8601, Japan

²Mitsubishi Tanabe Pharma Corporation, 1000, Shida-cho, Aoba-ku, Yokohama, 227-0033, Japan

[†]Authors contributed equally

Corresponding Author: Dr. Ryuji Kato, Graduate School of Pharmaceutical Sciences, Nagoya University, Furocho, Chikusa-ku, Aichi, 464-8601, Japan.

Tel.: +81-52-747-6811

Fax: +81-52-747-6813

E-mail: kato-r@ps.nagoya-u.ac.jp

Short title

Image-based prediction of myotube formation

Abstract

The development of new drugs depends on the efficiency of drug screening. Phenotype-based screening has attracted interest due to its considerable potency for the discovery of first-in-class drugs. In general, fluorescently-labeled imagery is the leading technique for phenotype-based screening; however, there are growing requirements to understand total culture profiles, which are unclear after end-point assays. In this study, we demonstrate that morphology-based cellular evaluation of unlabelled cells is an efficient approach to evaluate myotube formation assays. One of our aims was to study the myogenic differentiation process in C2C12 cells to discern the differences between cellular responses to different medium conditions (serum concentrations and insulin dosages). Our results show that predictive morphological profiles that strongly correlate with myogenic differentiation can be generated from myotube images, even in the confluent stage. The differentiation rate after 14 days can be quantitatively predicted with the highest accuracy by means of images taken on days 0–11.5. In addition, for the application of our morphology-based cellular evaluation, the effect of cyclic guanosine monophosphate (cGMP) on myogenic differentiation was analysed. Our results show that the quantitated morphological profile from these images can be an effective descriptor for analysis of the myotube-recovering effect of cGMP.

Keywords

Cellular morphology, morphology-based analysis, myoblast differentiation, cGMP

Introduction

Efficient drug screening is vital to the development of novel drugs. There are two categories of screening methods, target-based screening and phenotype-based screening. Target-based screening is a method of detecting drugs with high affinity by calculating interactions between candidate compounds and target proteins related to a symptom. As molecular biology and genetic engineering have advanced, target-based screening has achieved significant successes ¹. Nevertheless, phenotype-based screening has also attracted substantial interest. Phenotype-based screening is a method for identifying drugs with high specificity for phenotypic traits such as cell proliferation, cell morphology change, or cell death. This method has the advantage of enabling discovery of a drug candidate without the knowledge of its mechanism of action and has made useful contributions to finding first-in-class drugs ².

There are, however, several challenges posed by the cell-based assays performed during phenotype-based screening procedures. What limits the throughput of phenotypic screens is mostly the large number and complexity of the steps in the assay protocol (including washes), the slow speed of data acquisition and analysis, and the large amount of generated data that makes processing and analysis complicated and slow. In many cases, the assay windows are small and can reduce assay robustness and make screening a challenge when researchers use non-clonal cell lines or primary or iPSC-derived cells.

Another aspect of phenotype-based screening for which improvements are sought is the method of

measuring cellular responses. High-content screening (HCS) is a progressing technology enhanced by automation, which is currently leading such cell-based assays ^{3,4,5}. Data acquired by HCS are based on images of fluorescently labelled cells, in which multiple parameters describing cellular responses are analysed using image-processing techniques. For fluorescent cellular analysis, recent advances in technology relating to antibodies, fluorescent probes, and genetically introduced fluorescent proteins play important roles. However, the labelling-based schemes used in HCS are normally costly and time consuming, and their evaluation tends to be based only on an end-point assay.

As an alternative scheme for measuring cellular responses in cell-based assays, we have been investigating morphology-based cellular evaluation. This uses phase contrast microscopic images, which are processed by image- and data-processing techniques to extract multiple parameters of cellular morphology simultaneously. Advantages to this approach are (1) low-cost, (2) high-throughput, and (3) real-time detection and evaluation. We have reported the efficient application of this method to the quality assessment of mesenchymal stem cells ^{6,7} and induced pluripotent stem cells ⁸ and to the screening of PC12 cells ⁹.

In this work, we demonstrate that our morphology-based cellular evaluation method can be applied to myotube formation assays for drug screening. As a model case, we have evaluated the cellular responses of C2C12 cells, which form multi-nuclei cells and acquire morphological features of differentiating muscle cells ^{10,11}. Recently, such myotube formation assays have been recognized as

effective *in vitro* for the study of skeletal muscle pathologies including atrophy¹², dystrophy¹³, and sarcopenia¹⁴. The evaluation of myoblast differentiation can serve as an important screening tool, particularly for the study of myotube responses *in vivo*. In spite of the wide usage of myotube formation assays, they elicit several concerns. First, a differentiation assay takes a significant amount of time (1–2 weeks). Second, observation of differentiation can be naïve, because bulk measurement using stains such as myosin or myogenin¹⁵ suffers from a lack of sensitivity owing to heterogeneity in myotube formation, and can fail to detect the responses of small subpopulations. Third, counts of fused and elongating myotubes with multiple nuclei, which are some of the criteria of differentiated cells, can be biased. This is because there is no clear quantitative definition of such ‘fused and elongating myotubes’.

To overcome these concerns, we here show that image-based parameter description of myotubes can support the quantitative evaluation and prediction of myogenic differentiation. This involves measurement of C2C12 myotube differentiation responses under different medium conditions (serum concentrations and insulin doses), screening and quantitative analysis (Fig. 1). Our results show that images of confluent myotubes contain predictive information that strongly correlates with their actual differentiation state. As a result, the myotube differentiation rate after 14 days can be quantitatively predicted with the highest accuracy from images taken on days 0–11.5 without staining. Therefore, our scheme can be applied to reduce biases in myotube formation assays.

Materials & Methods

Cells and Cell Culture

C2C12 mouse myoblast cells were obtained from RIKEN Cell Bank (Tsukuba, Japan). C2C12 cells were seeded at a concentration of 5.0×10^4 cells in type I collagen-coated 24-well plates (AGC Techno Glass Co., Ltd., Shizuoka, Japan). C2C12 cells (passage 5) were maintained in the growth medium (Dulbecco's modified Eagle's medium [DMEM]; Wako Pure Chemical Industries, Ltd., Osaka, Japan) supplemented with 10% of foetal bovine serum (Life Technologies Japan, Ltd., Tokyo, Japan) and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (Nakarai Tesque Inc., Kyoto, Japan). When cells reached sub-confluent status, myogenic differentiation was induced by changing the medium to the differentiation medium [low-glucose DMEM (Wako Pure Chemical Industries, Ltd.) supplemented with 2% of horse serum (Life Technologies Japan, Ltd.) and 10 mM HEPES buffer] with or without insulin (1 $\mu\text{g}/\text{mL}$) (Wako Pure Chemical Industries, Ltd.) or 8-(4-chlorophenylthio)-guanosine 3',5'-cyclic monophosphate sodium salt (8-pCPT-cGMP) (0.1–1 mg/mL) (Sigma-Aldrich Japan Co., Tokyo, Japan) for a maximum of 20 days. The medium was refreshed every other day. For the 8-pCPT-cGMP effect assay, we evaluated three lots (passages 5, 8, and 11) of C2C12 cells with different passage numbers for obtaining a robust cellular effect.

Immuno-fluorescence staining

On days 2, 6, 10, 14, 18, and 22 after differentiation induction, cells were washed with

phosphate-buffered saline (PBS) (Nissui Pharmaceutical Co., Ltd., Tokyo, Japan) and fixed for 15 minutes in 2% paraformaldehyde (Wako Pure Chemical Industries, Ltd., Osaka, Japan). The cells were washed with PBS containing 0.2 M glycine (Wako Pure Chemical Industries, Ltd.) and permeabilised in PBS containing 0.3% Triton X-100 (Wako Pure Chemical Industries, Ltd.) for 30 minutes. After blocking in PBS containing 3% of bovine serum albumin (BSA) (Wako Pure Chemical Industries, Ltd.) overnight at 4°C, a primary antibody (anti-myosin heavy chain [MHC] mouse IgG, Merck Millipore, Billerica, MA, USA) was incubated for one hour at room temperature with the cell samples, and then a secondary antibody (Alexa Fluor 488–conjugated goat anti-mouse IgG (H+L) antibody, Cell Signaling Technology Inc., Danvers, MA, USA) was incubated for 30 minutes at room temperature. After immuno-staining, nuclei were stained with SYTOX Blue (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Image acquisition

Phase contrast and fluorescent images of C2C12 cells were acquired using a Biostation CT (Nikon Corp., Tokyo, Japan) at 4× magnification. One visual field in the centre position was analysed in each well with automatic focusing, and all phase contrast images were acquired once every six hours for a total of 22 days.

Image processing

All microscopic images were processed and quantified using CL-Quant software version 3.2 (Nikon Corp.). Phase contrast images were recognised in four steps: (1) background adjustment, (2) cell recognition, (3) filling in blank of cells, (4) removing noise, (5) eroding and dilating cells, and (6) quantification of cell morphology (Supplementary Information Fig. S1). Morphological profiles using six parameters (area, compactness, perimeter, length, width, and length/width [l/w] ratio) were determined in the recognised cellular objects after step (5).

Statistical analysis

Hierarchical clustering was applied to analyse the morphological profile in each culture condition (Conditions A, B, C, and D). The averages (AVE) and standard deviations (SD) of morphological profiles were calculated from data on 500–1000 cells per image at each time point (Supplementary Information Fig. S1). The calculated raw data were then normalised for each parameter by standard normalisation (the normalisation procedure set the mean of the data to 0 and standard deviation to 1.0). Therefore, the parameters finally used in this work had no units and were represented by the data whose centre was 0 and which were spread toward both positive and negative values, with their standard deviation being 1.0. Hierarchical clustering was performed with R software (version 3.1.0, R Development Core Team, <https://www.r-project.org/>) using a complete linkage by correlation coefficient.

Construction of a prediction model

Total morphological profiles obtained from phase contrast images were used for construction of regression models that predict the MHC-positive cellular area. In this work, least absolute shrinkage and selection operator (LASSO) regression models trained with input data (morphological profile) were used in combination with output data (the differentiation rate experimentally determined by MHC staining and quantification of the stained area). Via construction of the LASSO regression model, only the best combinations of morphological parameters were automatically explored for the prediction (Table 1). Image data used for morphological profile extraction were taken from the non-induction growth period (0–2 days after seeding) and the myogenic differentiation period (2–22 days after seeding; total 20 days). The prediction model was constructed by training 20 datasets (condition A: 4 datasets; condition B: 12 datasets; condition C: 4 datasets). Each dataset was produced by end-point MHC staining; therefore, four types of prediction model were constructed: on day 10 (0–10 days of morphology and MHC staining on day 10), day 14 (0–14 days of morphology and MHC staining on day 14), day 18 (0–18 days of morphology and MHC staining on day 18), and day 22 (0–22 days of morphology and MHC staining on day 22). In addition, to investigate the effect of time course data usage on prediction accuracy, total time course information was shortened by each time point, and model performance was examined comprehensively. In such time course data reduction examination, prediction accuracy was plotted by rooted mean square error (RMSE), which is derived from differences between ‘predicted value’ and ‘experimentally determined value’. Simply

put, the lower the RMSE value, the higher is the prediction performance. Each prediction model's performance was evaluated by leave-one-out cross-validation. R (version 3.1.0, R Development Core Team, <https://www.r-project.org/>) was used in the analysis.

Results

Measurement of myotube differentiation in different culture conditions

It is well established that a reduction of serum concentration in a culture medium triggers differentiation of C2C12 cells into myotubes^{16,17}. As part of this process, myoblasts start to fuse to each other and elongate by escaping from the cell cycle in favour of proliferation. As a result, large and elongated cells with multiple nuclei, a measure of myotube formation that correlates with the MHC, start to appear in the population. To establish a quantitative image-based measurement scheme for describing such myotube formation, we first prepared three types of culture conditions: a negative control consisting of the differentiation medium containing 10% of foetal bovine serum (FBS), which is commonly used as a growth medium (condition A), the differentiation medium containing 2% of horse serum (condition B), and the differentiation medium containing 2% of horse serum and 1 µg/mL insulin (condition C). Insulin is known to promote muscle differentiation of C2C12 cells^{18,19}. During 22 days of differentiation culture, phase contrast images were acquired as 'label-free morphology data' at 6 h intervals. In parallel with this culture, differentiation rates were experimentally determined by MHC staining (on days 2, 6, 10, 14, 18, and 22) (Fig. 2A).

In the stained images, if differentiated myotubes are defined as ‘large cells with multiple nuclei resulting from cellular fusion’, the C2C12 cell line showed such cells in conditions B and C, where their differentiation was incomplete in condition A, resulting in elongating cells that remained thin and single. In conditions B and C, differentiated myotubes started to appear from day 10, however their growth gradually stopped or decreased. By comparison, in condition A, myoblasts did elongate, but did not show fused large differentiated myotubes throughout the culture.

Conventionally, such differentiated myotubes are counted in stained images at researchers’ discretion, and there is no quantitative definition. In our analysis, stained myotubes with MHC over $4000 \mu\text{m}^2$ were defined as ‘differentiated myotubes’ (Fig. 2B), and their average number (Fig. 2C) and size distribution (Fig. 2D) were analysed. These results show that our protocol consists of clear condition-dependent differences that can be clarified by fluorescent image analysis.

Quantitative prediction of myotube differentiation by morphological analysis of phase contrast images

To investigate whether the morphological profile of non-labelled cells can represent the effects of different culture conditions, we performed morphology-based evaluation. Briefly, we first identified the cellular objects in images of each time point: not only the differentiated large myotubes but also the undifferentiated cells in the visual field. Then, we measured six morphological parameters (area, compactness, perimeter, length, width, and the l/w ratio) for every cellular object in the image. These

were converted into statistics representing the ‘group of cells’ in the image; averages (AVE) and standard deviations (SD) of six morphological parameters calculated from data on 500–1000 cells per image (see Materials and Methods). This morphological profile was determined at 56 time points (6 h intervals) and was used as the total morphological profile of a condition (672 parameters = 6 parameters \times 2 statistics \times 56 time points) (Fig. 3A). To compare these profiles, hierarchical clustering, which objectively describes correlations between multi-parameter profiles, was applied. The results for each culture condition condensed as individual clusters, indicating that the morphological profiles of myotubes are similar, and represent the differences between culture conditions.

Next, to investigate predictive performance, the prediction model was constructed by training the data sets of morphological parameters from the phase contrast images (0–14 days), and MHC-positive cellular area determined by means of images of fluorescently stained cells (day 14) (Fig. 3B). As a result, the prediction of the MHC-positive cellular area, which can be obtained only destructively after 14 days, can be predicted with images from day 0–4, which start to manifest low RMSE. The prediction accuracy increases more when the images being analysed covered more than 0–10 days and is maximised with images from 0–11.5 days (accuracy = 95%, based on 19/20 images; Fig. 3C). Such a positive effect of accumulation of the morphological information for its high prediction performance can only be obtained because our evaluation is live and continuous and involves non-destructive image assessment. Considering the possibility of use in the screening, this

non-destructive assay should also be advantageous because of its low cost and high throughput. Highly weighted parameters selected for the prediction model were length, width, and l/w ratio, representing the 'elongation feature' (Table 1). In contrast, other parameters relating only to 'size' were not selected. From the time course, both the early and late periods of differentiation were selected. From these results, the prediction rule could be interpreted for further understanding of the phenotypes.

Application of morphological profiles of phase contrast images for estimating the performance of a new culture condition

To extend the usage of the morphological profiles that we extracted from continuous images of C2C12 cells during their differentiation, we examined them in an additional culture condition (condition D). Condition D included cGMP (0.1 mg/mL) from the first day. cGMP has been reported to reverse mitochondrial dysfunction in C2C12-derived myotubes^{20,21}, but its performance and usage protocol have not been quantitatively examined. First, the effects of condition D were monitored via phase contrast images and MHC staining patterns (Fig. 4A). The MHC staining results gradually increased during the culture period and reached a maximum on day 14. The quantification of the fluorescent images provided quantitative kinetics of differentiation and correlated to observation (Fig. 4B, 4C). Secondly, by using the total morphological profile representing differentiation on day 14, condition D was clustered together with conditions A, B, and C (Fig. 5). As a result, condition D

clustered in the same cluster as condition B (2% of horse serum) and diverged from condition A (10% of FBS) and condition C (2% of horse serum with insulin). Thirdly, differentiation was predicted in condition D using the constructed prediction model, which predicts the MHC cellular area after 14 days of differentiation. The prediction accuracy was 100% (4 images / 4 images).

Discussion

To investigate the applicability of morphological analysis of non-stained images, we applied morphological evaluation to a myotube formation assay. Myotube formation assays are morphology-based assays frequently used in the study of muscle-related symptoms. Our objective was to investigate the effectiveness of our profiling method, which extracts morphological profiles representing time course changes of cellular populations for predictive evaluation of myoblast differentiation.

In this work, we compared four different culture conditions using morphological profiles extracted from phase contrast images. Using over 500 morphological parameters, the similarities between culture conditions were clustered without training their labels. In other words, our results indicate that time course morphological changes can reflect the characteristics of myoblast responses to their culture media. Given that our method requires only label-free phase contrast images, such live monitoring information can be obtained at low cost and in a high-throughput manner by the phenotype-based screening.

To further explore the potential of such ‘morphological profiles’, we then trained a computer to construct a quantitative prediction model for non-invasive evaluation of cellular differentiation. Prediction of day 14 MHC cellular staining state showed that the time course image usage can predict the future differentiation rate solely from a morphological profile.

One of the reasons we were able to attain such high predictive performance was our method’s characteristic morphological feature extraction. In our profiling, image data are converted into parameters, and once measured, individual morphological parameters are finally converted into statistical descriptors of a cellular population (500–1000 cells in total). Previously, we reported the effectiveness of such profiling parameters in mesenchymal stem cells. On the basis of the data in this work, we also showed that morphological parameters are descriptive when they represent the ‘heterogeneity’ of a cellular population. Especially in myotube formation, not all the cells in a population respond to the myotube, and therefore the description parameter for the balance of a cellular population can be effective. Moreover, our previous reports also indicate that quantification of time course changes is also effective for analysis. After 14 days of morphological transition, the clustering and prediction models showed high performance. The weight interpretation in our model could also provide insight into the phenotypic reactions of myoblasts. In the weighted parameter list, both AVE and SD, representing population characteristics – together with their substantial change rate from the early to late period – were found to be important morphological indicators.

To the prediction of MHC-positive cellular area after 14 days, we also coupled other time point

predictions in parallel: prediction of the MHC-positive cellular area on days 10, 18, and 22 (Supplementary Information Fig. S2). When the image data use period was comprehensively examined for exploring the best prediction combination for the MHC-positive area on day 10, no model could show significant prediction performance. Such failure of prediction model construction was also observed in the prediction of the MHC-positive area on days 18 and 22. This result indicates that not every type of morphological profile could lead to the prediction of myogenic differentiation. From the aspect of data set quality, at least two factors could trigger such failure of prediction. One is the absence of positive data, i.e. of positively stained differentiated myotubes, on day 10. At this stage, myoblasts have not yet been differentiated, therefore not enough positive data (stained images) were in the training dataset. In contrast, for differentiation culture beyond 14 days, there were MHC-stained myotubes, albeit thin and not greatly expanding, even in the visual field of the negative control (condition A). Around this stage, the positive controls (conditions B–D) lose stained large myotubes rapidly. In other words, after day 14, the ‘answer’ for training, in particular, the rate of staining of myotubes decreases in the total stained area, thus resulting in a drop of the signal-to-noise ratio. With such data, machine learning does not succeed in producing high-performance models. Therefore, this investigation suggests that morphological prediction models have to be carefully designed to obtain fair sample numbers of both positive and negative data that can reflect the assay’s objective.

As a test case for our morphological prediction scheme for the evaluation of culture conditions, a

cGMP addition effect was evaluated in condition D. Biologically, the MHC staining showed that GMP addition did not alter differentiation as compared with condition B. Condition D showed a morphological profile nearly identical to that of condition B in clustering analysis, and its MHC-positive cellular area prediction could be accurately performed using our model. This result indicated that such ‘non-toxic’ or ‘non-damaging’ evaluation can be estimated by morphological profiling, or prediction evaluation using machine learning models.

To further extend our evaluation of the effects of culture conditions on predictive performance, we next added cGMP in condition B, only on day 18, the point at which MHC-stained cellular area decreased following its differentiation peak. The MHC staining result showed that 10×, but not 1×, cGMP addition significantly recovered the MHC staining cellular area (Supplementary Information Fig. S3). Since this recovery occurred in six days, the previously constructed prediction model trained with 14 days of morphological profiles could not be applied. However, when the morphological parameters were applied to show a time course change, a characteristic change reflecting the effects of 10× cGMP promoting muscle differentiation could be visualised. Although the sensitively changed morphological parameters were different from the weighted parameters in the prediction model (0–14 days), detailed changes occurring in response to cGMP could be detected. Therefore, such morphological profiling can be considered applicable to such compound screening for differentiation culture.

Moreover, because of its non-invasive style, our morphological profiling can be combined with

other types of target-based assays apart from myotube differentiation assays. Non-invasive phenotypic profiling can enhance other drug screening assays, which have required conventional cell-staining assays aside from the target-based assays.

To further improve the prediction performance of our presented method, introduction of other image quantification techniques from advanced high-content image analysis is an effective approach. In our present analytical procedure, we simply measured the myotube differentiation rate via the MHC-positive area and used these data for machine learning as teacher signals. Nonetheless, there are other quantification concepts that have been proposed in the studies on HCS e.g. the fusion index reported by Young *et al.* ²², which can be a more stringent descriptor for a disease phenotype. To combine wider varieties of both multi-colour marker staining and the new concepts of the parameter index is certainly the aim of our next study, which is intended to enhance the performance of our proposed morphology-based analysis.

In summary, based on our evaluation method using morphological profiling of myoblast populations responding to myotube differentiation conditions, our data indicate the good potential of our morphological analysis concept for application to various cell differentiation assays. Our proposed concept can reduce the cost and can increase the throughput of cell-based screening procedures greatly because it requires only non-stained continuous images of cells. Inhibitors or enhancers found through cell differentiation studies are among leading therapeutic compounds, and our concept should facilitate and promote studies of muscle differentiation. It is a fact that whether

we can robustly apply our methodology to other types of myoblasts is the next important question to address in detail. Nevertheless, our morphological-analysis concept based on statistical features of groups of cells has been successfully applied to a variety of cell types (mesenchymal stem cells ^{6,7}, adrenal gland cells ⁹, and other cells [PCT Patent WO2011021391A1]) including different lots. Therefore, we consider the applicability of our methodology broad. Future work may include extension of our concept to a wider range of compound screening procedures involving patient-derived human myoblasts, which are now expected to be carried out in the studies on myotonic dystrophy type 1 ^{23,24}.

Acknowledgements

This research was funded by the Japan Science and Technology Agency (JST) Program for Creating STart-ups from Advanced Research and Technology (START Program) program.

References

1. Rix, U.; Superti-Furga, G. Target Profiling of Small Molecules by Chemical Proteomics. *Nat. Chem. Biol.* **2009**, *5*, 616–624.
2. Moffat, J. G.; Vincent, F.; Lee, J. A.; et al. Opportunities and Challenges in Phenotypic Drug Discovery: An Industry Perspective. *Nat. Rev. Drug Discov.* **2017**, *16*, 531–543.
3. Singh, S.; Carpenter, A. E.; Genovesio, A. Increasing the Content of High-Content Screening: An Overview. *J. Biomol. Screen.* **2014**, *19*, 640–650.
4. Kerz, M.; Folarin, A.; Meleckyte, R.; et al. A Novel Automated High-Content Analysis Workflow Capturing Cell Population Dynamics from Induced Pluripotent Stem Cell Live Imaging Data. *J. Biomol. Screen.* **2016**, *21*, 887–896.
5. Stengl, A.; Hörl, D.; Leonhardt, H.; et al. A Simple and Sensitive High-Content Assay for the Characterization of Antiproliferative Therapeutic Antibodies. *SLAS Discov.* **2017**, *22*, 309–315.
6. Matsuoka, F.; Takeuchi, I.; Agata, H.; et al. Morphology-Based Prediction of Osteogenic Differentiation Potential of Human Mesenchymal Stem Cells. *PLoS One* **2013**, *8*, e55082.
7. Sasaki, H.; Takeuchi, I.; Okada, M.; et al. Label-Free Morphology-Based Prediction of Multiple Differentiation Potentials of Human Mesenchymal Stem Cells for Early Evaluation of Intact Cells. *PLoS One* **2014**, *9*, e93952.
8. Nagasaka, R.; Matsumoto, M.; Okada, M.; et al. Visualization of Morphological Categories of

- Colonies for Monitoring of Effect on Induced Pluripotent Stem Cell Culture Status. *Regen. Ther.* **2017**, *6*, 41–51.
9. Kawai, S.; Sasaki, H.; Okada, N.; et al. Morphological Evaluation of Nonlabeled Cells to Detect Stimulation of Nerve Growth Factor Expression by Lyconadin B. *J. Biomol. Screen.* **2016**, *21*, 795–803.
 10. David, Y.; Ora, S.; Serial Passaging and Differentiation of Myogenic Cells Isolated from Dystrophic Mouse Muscle. *Nature* **1977**, *270*, 725–727.
 11. Girgis, C. M.; Clifton-Bligh, R. J.; Mokbel, N.; et al. Vitamin D Signaling Regulates Proliferation, Differentiation, and Myotube Size in C2C12 Skeletal Muscle Cells. *Endocrinology* **2014**, *155*, 347–357.
 12. Powers, S. K.; Lynch, G. S.; Murphy, K. T.; et al. Disease-Induced Skeletal Muscle Atrophy and Fatigue. *Med. Sci. Sports Exerc.* **2016**, *48*, 2307–2319.
 13. Fairclough, R. J.; Wood, M. J.; Davies, K. E. Therapy for Duchenne Muscular Dystrophy: Renewed Optimism from Genetic Approaches. *Nat. Rev. Genet.* **2013**, *14*, 373–378.
 14. Muscaritoli, M.; Anker, S. D.; Argilés, J.; et al. Consensus Definition of Sarcopenia, Cachexia and Pre-Cachexia: Joint Document Elaborated by Special Interest Groups (SIG) “Cachexia-Anorexia in Chronic Wasting Diseases” and “Nutrition in Geriatrics.” *Clin. Nutr.* **2010**, *29*, 154–159.
 15. Khurana, A.; Dey, C. S. Subtype Specific Roles of Mitogen Activated Protein Kinases in

- L6E9 Skeletal Muscle Cell Differentiation. *Mol. Cell. Biochem.* **2002**, 238, 27–39.
16. David, Y.; Ora, S. A Myogenic Cell Line with Altered Serum Requirements for Differentiation. *Differentiation* **1977**, 7, 159–166.
 17. Lawson, M. A.; Purslow, P. P. Differentiation of Myoblasts in Serum-Free Media: Effects of Modified Media Are Cell Line-Specific. *Cells Tissues Organs* **2000**, 167, 130–137.
 18. Florini, J. R.; Ewton, D.Z.; Coolican, S.A. Growth Hormone and the Insulin-like Growth Factor System in Myogenesis. *Endocr. Rev.* **1996**, 17, 481–517.
 19. Conejo, R.; Valverde, A. M.; Benito, M.; et al. Insulin Produces Myogenesis in C2C12 Myoblasts by Induction of NF- κ B and Downregulation of AP-1 Activities. *J. Cell. Physiol.* **2001**, 186, 82–94.
 20. Mitsuishi, M.; Miyashita, K.; Itoh, H. cGMP Rescues Mitochondrial Dysfunction Induced by Glucose and Insulin in Myocytes. *Biochem. Biophys. Res. Commun.* **2008**, 367, 840–845.
 21. Moro, C.; Lafontan, M. Natriuretic Peptides and cGMP Signaling Control of Energy Homeostasis. *AJP Hear. Circ. Physiol.* **2013**, 304, H358–H368.
 22. Young, J.; Margaron, Y.; Fernandes, M.; et al. MyoScreen, a High-Throughput Phenotypic Screening Platform Enabling Muscle Drug Discovery. *SLAS Discov.* **2018**, doi: 10.1177/2472555218761102.
 23. Furling, D.; Doucet, G.; Langlois, M. A.; et al. Viral Vector Producing Antisense RNA Restores Myotonic Dystrophy Myoblast Functions. *Gene Ther.* **2003**, 10, 795–802.

24. Matloka, M.; Klein, A. F.; Rau, F.; et al. Cells of Matter — In Vitro Models for Myotonic Dystrophy. *Front Neurol.* **2018**, *9*, 1–9.

Table 1.

Time [day]	Parameter	AVE	SD	Coefficient ($\times 10^3$)
0.5	width		○	-1.024
1.75	l/w ratio	○		8.828
1.75	l/w ratio		○	25.417
3	width		○	3.026
3.5	length		○	-2.010
4.75	l/w ratio	○		-0.998
6.5	l/w ratio		○	-4.640
6.75	l/w ratio	○		-2.490
7.5	l/w ratio	○		-9.057
8.25	compactness	○		-1.977
8.25	perimeter	○		-17.213
9.25	length		○	69.648
10.75	width		○	-34.169

1 **Table Legend**

2 Table 1. Contribution of selected morphological parameters to the LASSO prediction model
3 between days 0 and 11.5.

4

5 **Figure Legends**

6 Figure 1. The experimental scheme used in this study. Sub-confluent C2C12 cells were cultured in
7 three different conditions (A, B, and C). On days 2, 6, 10, 14, 18, and 22, the cells were stained for
8 myosin heavy chain (MHC), and its staining area was quantified by image analysis. During culture,
9 the acquired images were processed to extract cellular morphology using six morphological
10 parameters. Using the morphological parameters and the MHC-stained area as a training dataset, a
11 prediction model was constructed.

12

13 Figure 2. Measurement of myogenic differentiation. (A) Representative phase contrast images and
14 immunofluorescent images of C2C12 cells stained for MHC (green) and SYTOX Blue (blue). The
15 cells were maintained in three different conditions (A, B, and C), which differ in serum
16 concentration and insulin dose. Scale bar, 400 μm . (B) Measurement of myotubes. Images of
17 fluorescently MHC-labelled cells were binarised, and each cellular object was recognised. In the
18 cellular population, objects that exceed 4000 μm^2 were defined as ‘myotubes’ in our analysis. (C)
19 Transition of an MHC-positive cellular area in response to culture conditions. Error bars indicate

1 means \pm SD for four independent images. (D) Transition of myotube population in response to
2 culture conditions. Three solid lines on the bee swarm plot indicate the first, second (median), and
3 third quantiles from the above.

4

5 Figure 3. Morphological analysis of myotube formation in response to culture conditions. (A)
6 Hierarchical clustering of morphological profiles obtained from phase contrast images within
7 growth induction for two days and growth induction between days 0 and 14. Conditions A (N = 4),
8 B (N=12), and C (N=4) are indicated by capital letters. The dashed line on the hierarchical tree
9 indicates the point of no correlation (correlation coefficient = 0). (B) Transition of performance on
10 MHC-positive area prediction between days 0 and 14. The dots indicate the root mean square error
11 (RMSE) value as a function of time using different periods of a morphological profile. For example,
12 the blank dot indicates the RMSE of the prediction model trained with 0–6 hours of images. X-axis:
13 Length of time for determining a morphological profile for the prediction; Y-axis: RMSE value. The
14 red arrow indicates the minimum RMSE point (image usage = 0–11.5 days), prediction performance
15 of which is shown in Fig. 3C. The dotted line indicates the predictive performance of the NULL
16 model, which is the negative control. (C) Performance on MHC-positive area prediction at 14 days
17 by morphological profile (days 0–11.5). X-axis: the experimentally determined value, Y-axis: the
18 computationally predicted value only from the morphological profile. Condition A (red), condition
19 B (blue), and condition C (green). The grey region covering the dots indicates the possible deviation

1 area that can be affected by the experimental error (defined as \pm SD of experimental data). If the
2 dots are in the grey region, the prediction has higher reliability.

3

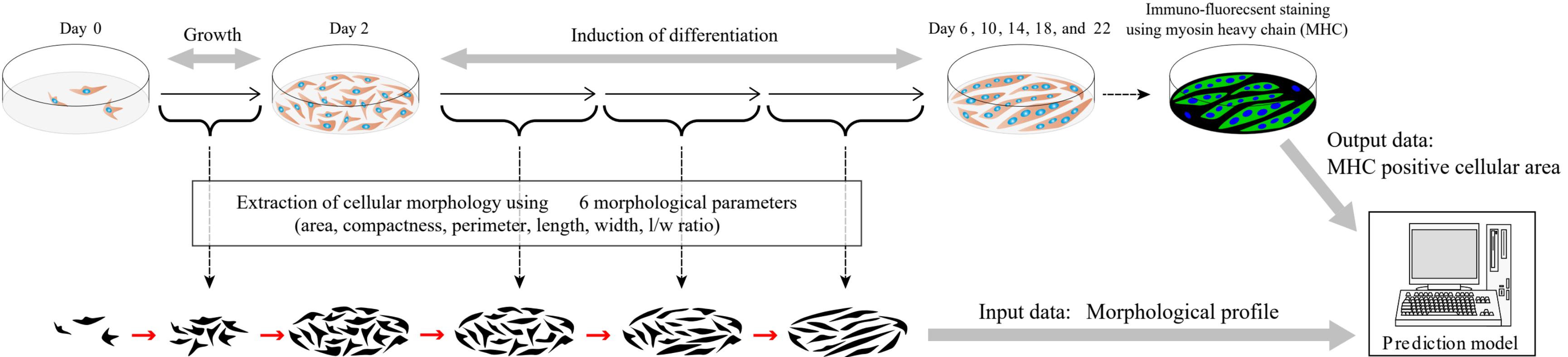
4 Figure 4. Measurement of myogenic differentiation in condition D. (A) Representative phase
5 contrast images and immunofluorescent images of C2C12 cells stained for MHC (green) and
6 SYTOX Blue (blue). Scale bar, 400 μ m. (C) Transition of the MHC-positive cellular area in
7 response to the culture conditions. Error bars indicate means \pm SD for four independent images. (D)
8 Transition of myotube population in response to the culture conditions. The three solid lines on the
9 bee swarm plot indicate the first, second (median), and third quantiles from above.

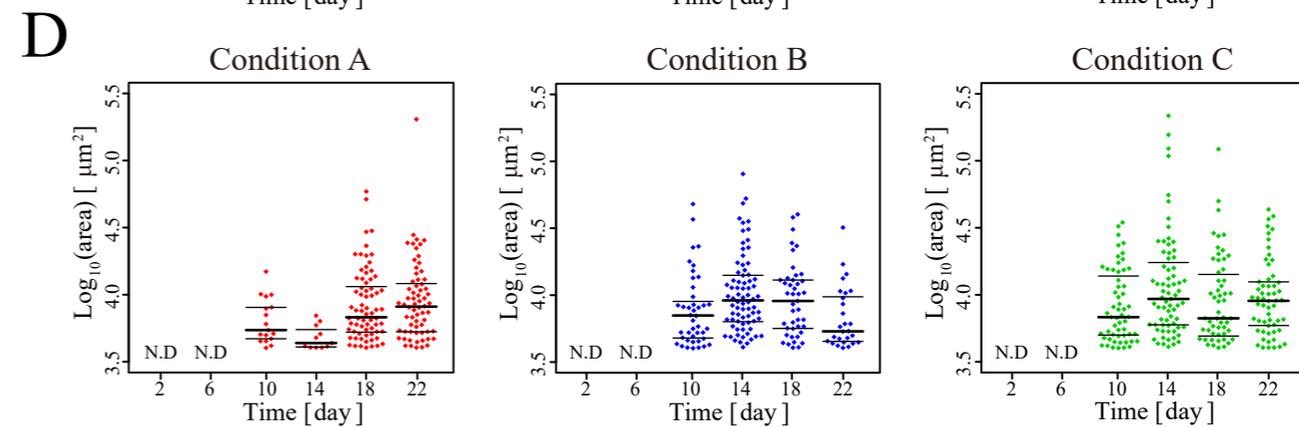
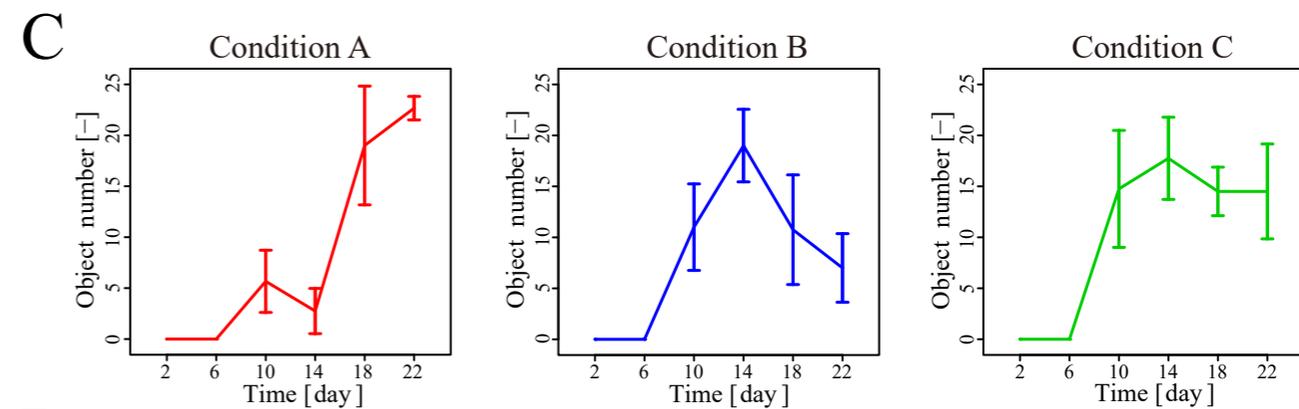
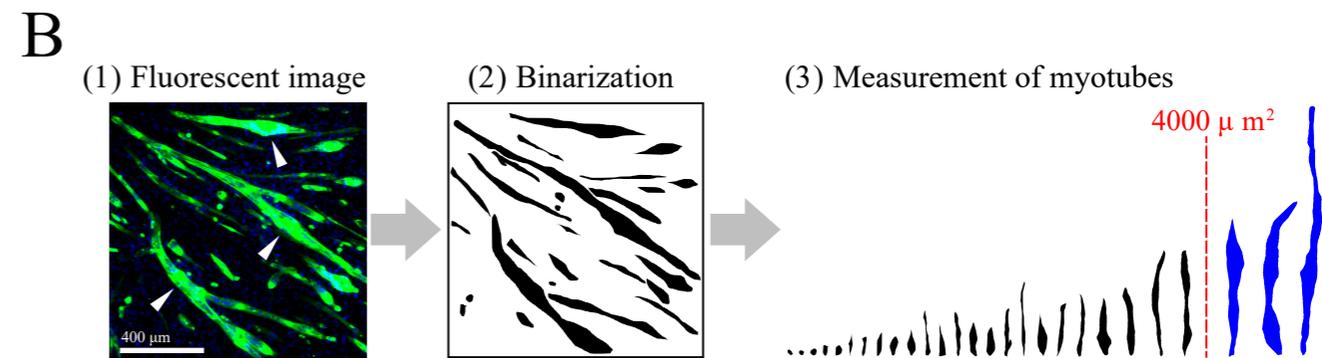
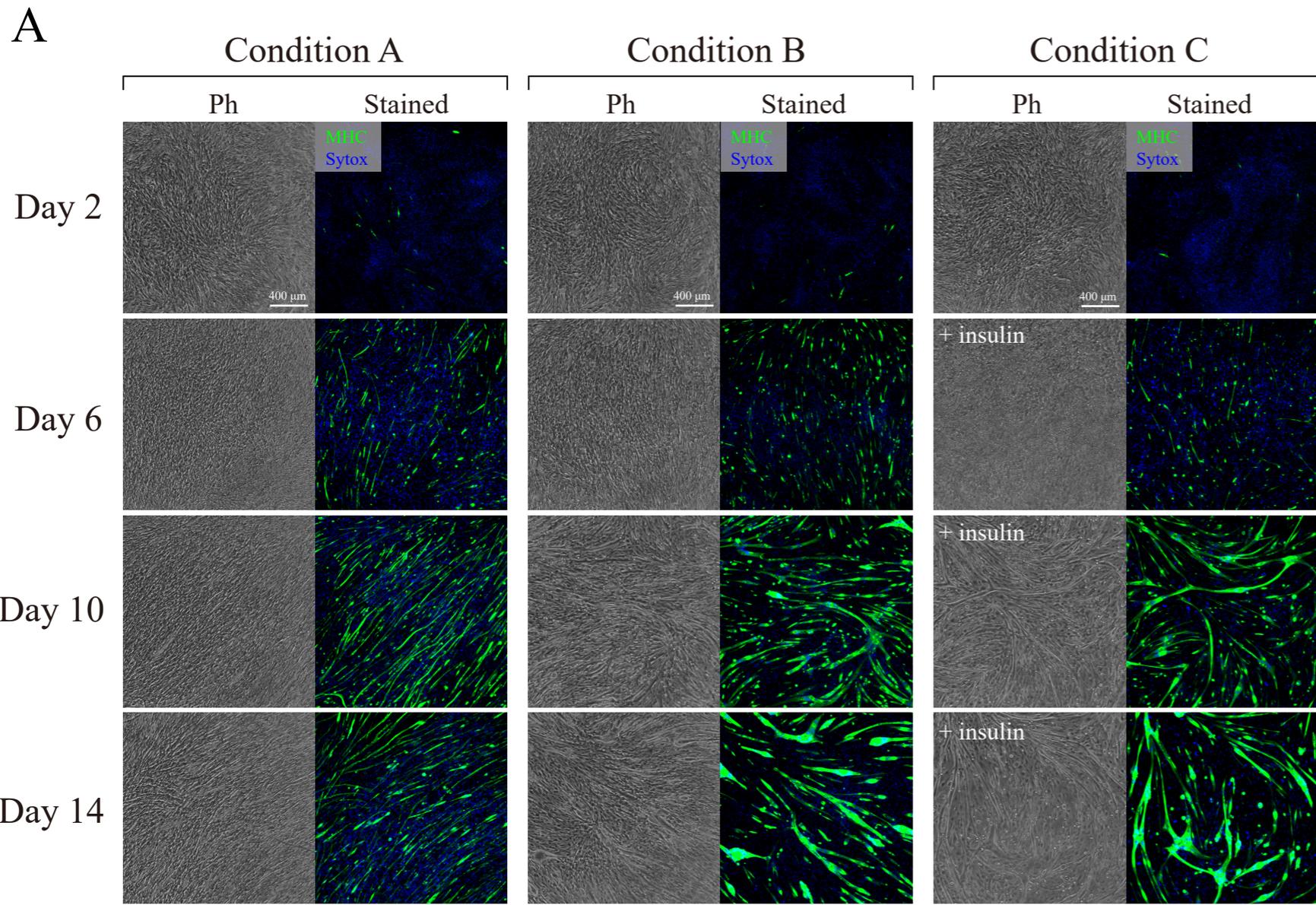
10

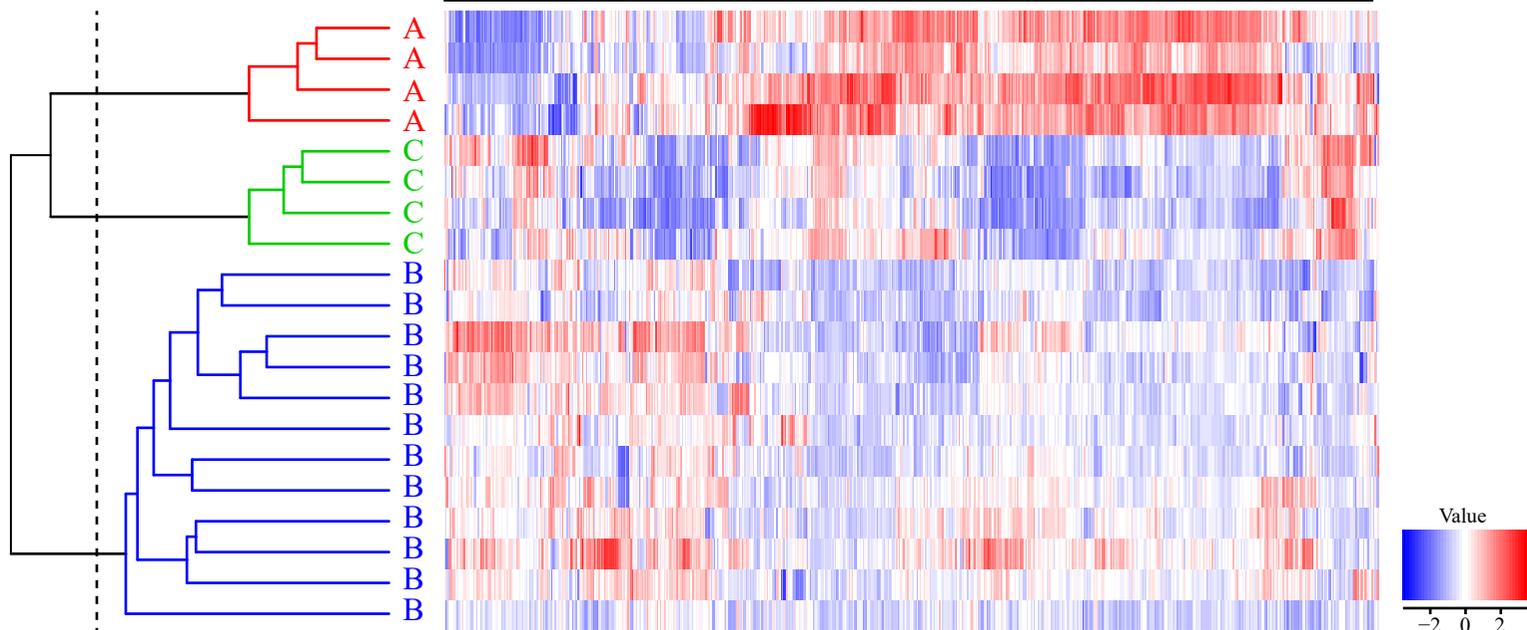
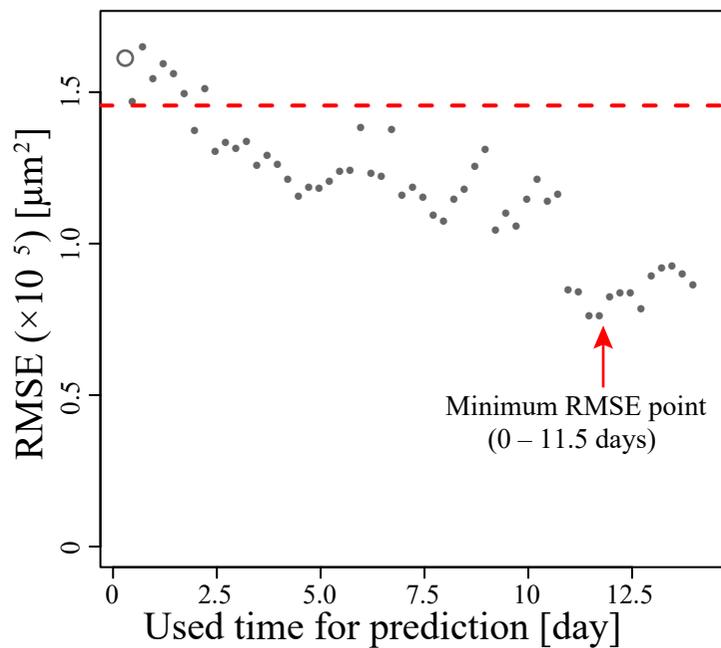
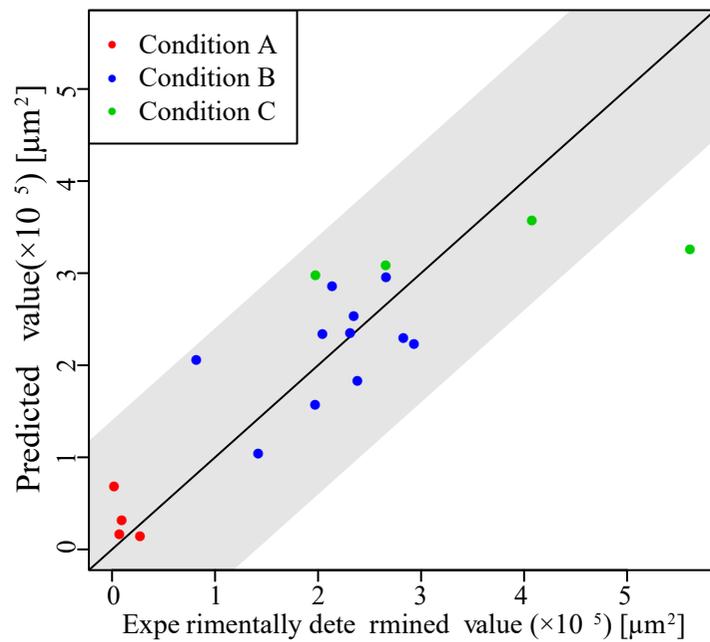
11 Figure 5. Hierarchical clustering of morphological profiles obtained from phase contrast images
12 within 0–14 days. Condition A (N=4), B (N=12), C (N=4), and D (N=4) are indicated by capital
13 letters. The dashed line on the hierarchical tree indicates the point of no correlation (correlation
14 coefficient = 0).

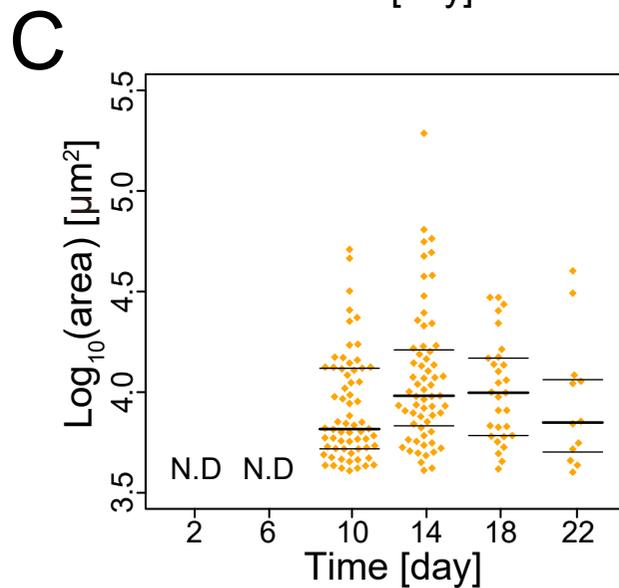
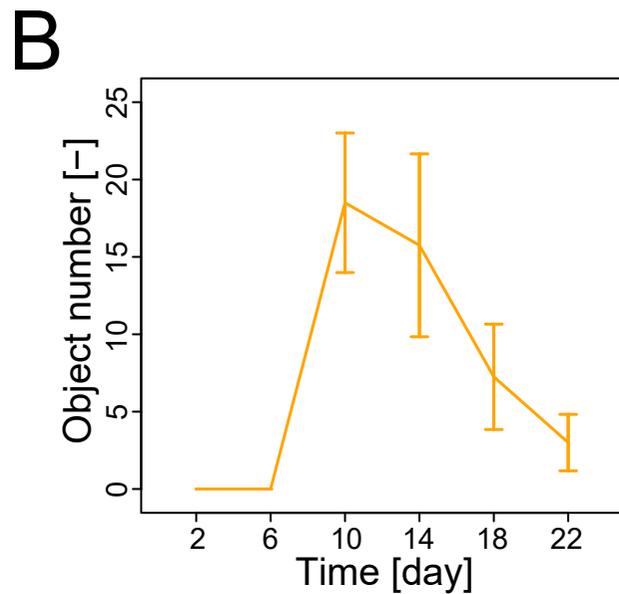
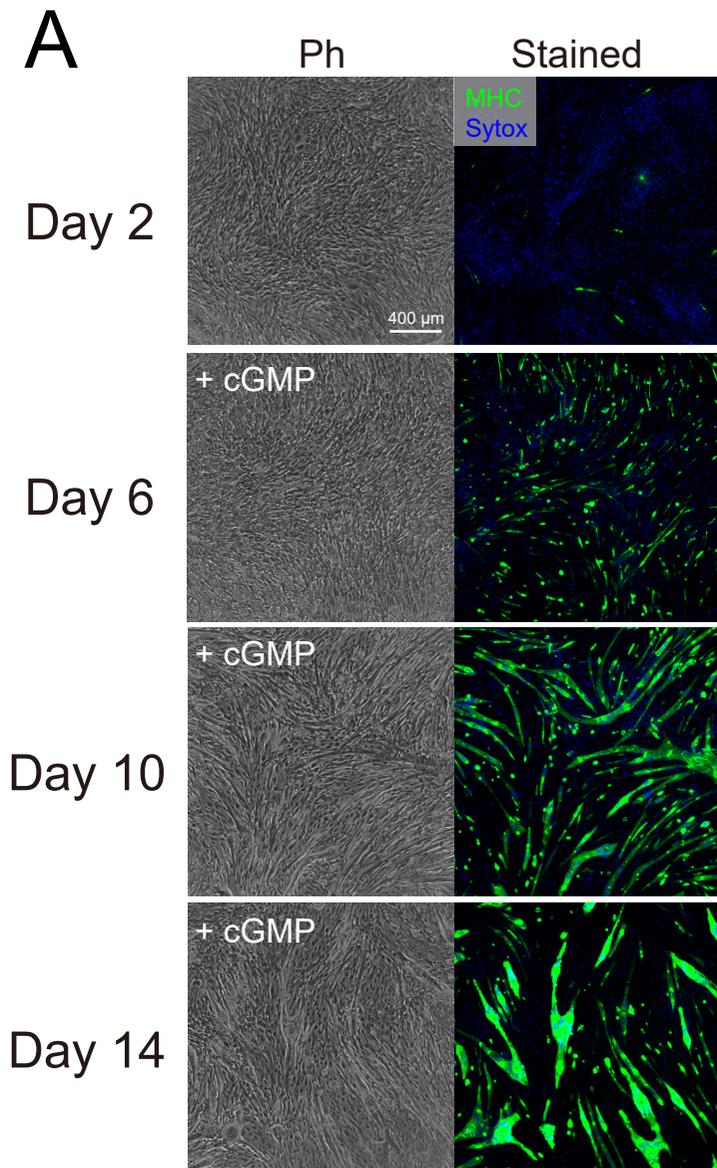
Image acquisition (6-h interval) for morphology-based analysis

Endpoint assay for experimental data



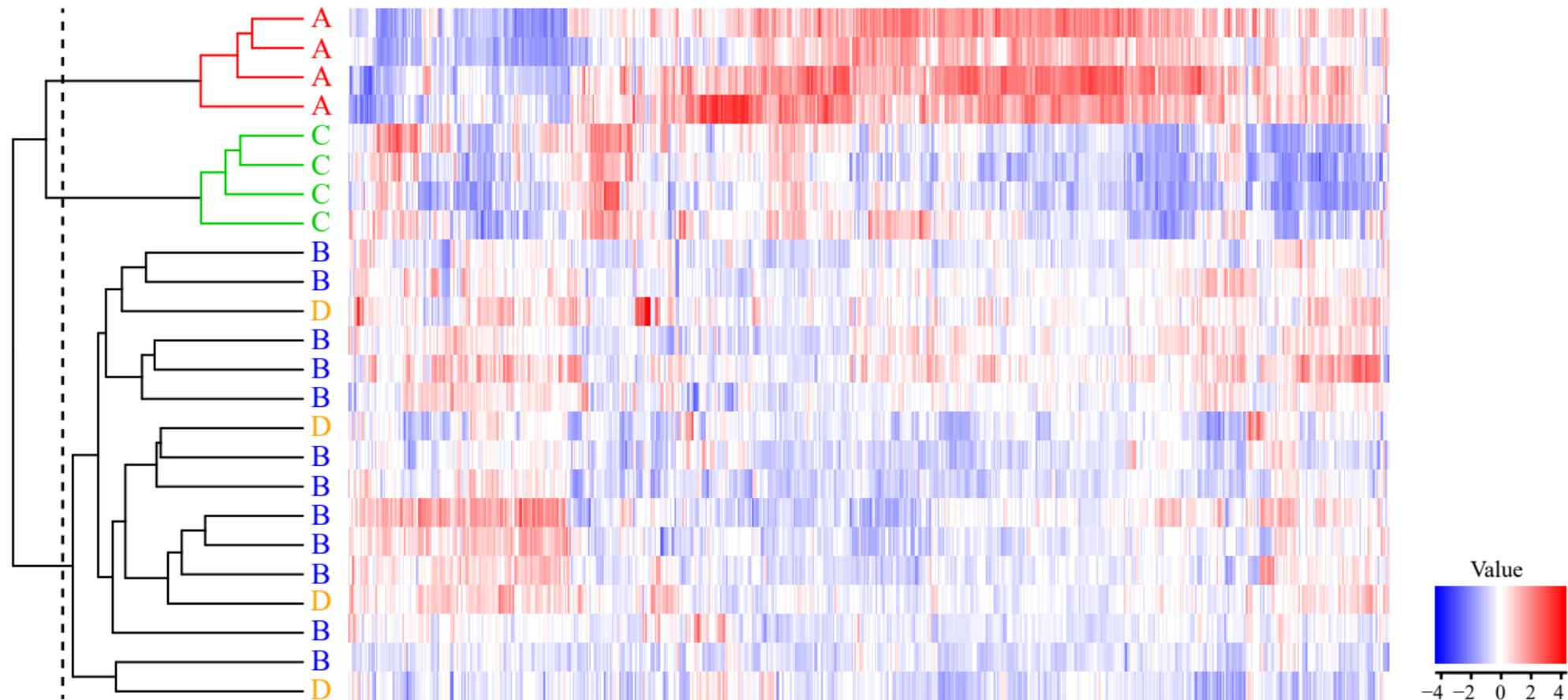


A672 parameters = 6 Morphological parameters \times 2 statistics \times 56 time-points**B****C**



672 parameters = 6 Morphological parameters \times 2 statistics \times 56

time-points



Morphology-based analysis of myoblasts for prediction of myotube formation

Kiyoshi Ishikawa^{1,2†}, Kei Yoshida^{1†}, Kei Kanie¹, Kenji Omori², Ryuji Kato¹

¹Graduate School of Pharmaceutical Sciences, Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-8601, Japan

²Mitsubishi Tanabe Pharma Corporation, 1000, Shida-cho, Aoba-ku, Yokohama, 227-0033, Japan

[†]Authors contributed equally

Corresponding Author: Dr. Ryuji Kato, Graduate School of Pharmaceutical Sciences, Nagoya University, Furocho, Chikusa-ku, Aichi, 464-8601, Japan.

Tel.: +81-52-747-6811

Fax: +81-52-747-6813

E-mail: kato-r@ps.nagoya-u.ac.jp

Keywords

Cellular morphology, morphology-based analysis, myoblast differentiation, cGMP

Supplementary Information

Figure Legends

Figure S1. The scheme of morphological analysis of confluent myotube cells.

Starting with the phase contrast image of myotube confluence status, the image was next processed through five filters to grasp recognisable objects that reflect the status of the myotubes. In this processing procedure, the ultimate goal was to extract the phenotypic morphological character of the confluent cellular status, rather than to segment the details of objects. Because there is no clear definition of myotube size, our approach was to recognise the candidate cellular objects, which may include all stages of myoblasts differentiating toward the myotube, and to statistically represent the morphological feature of a 'group of cells in the image' by means of the average and standard deviation of six individual morphological parameters. By accumulating these parameters throughout the time course of the study, our morphological parameters represent not only each stage of the morphological pattern but also the time course transitions of morphological patterns of 'differentiating myoblasts'.

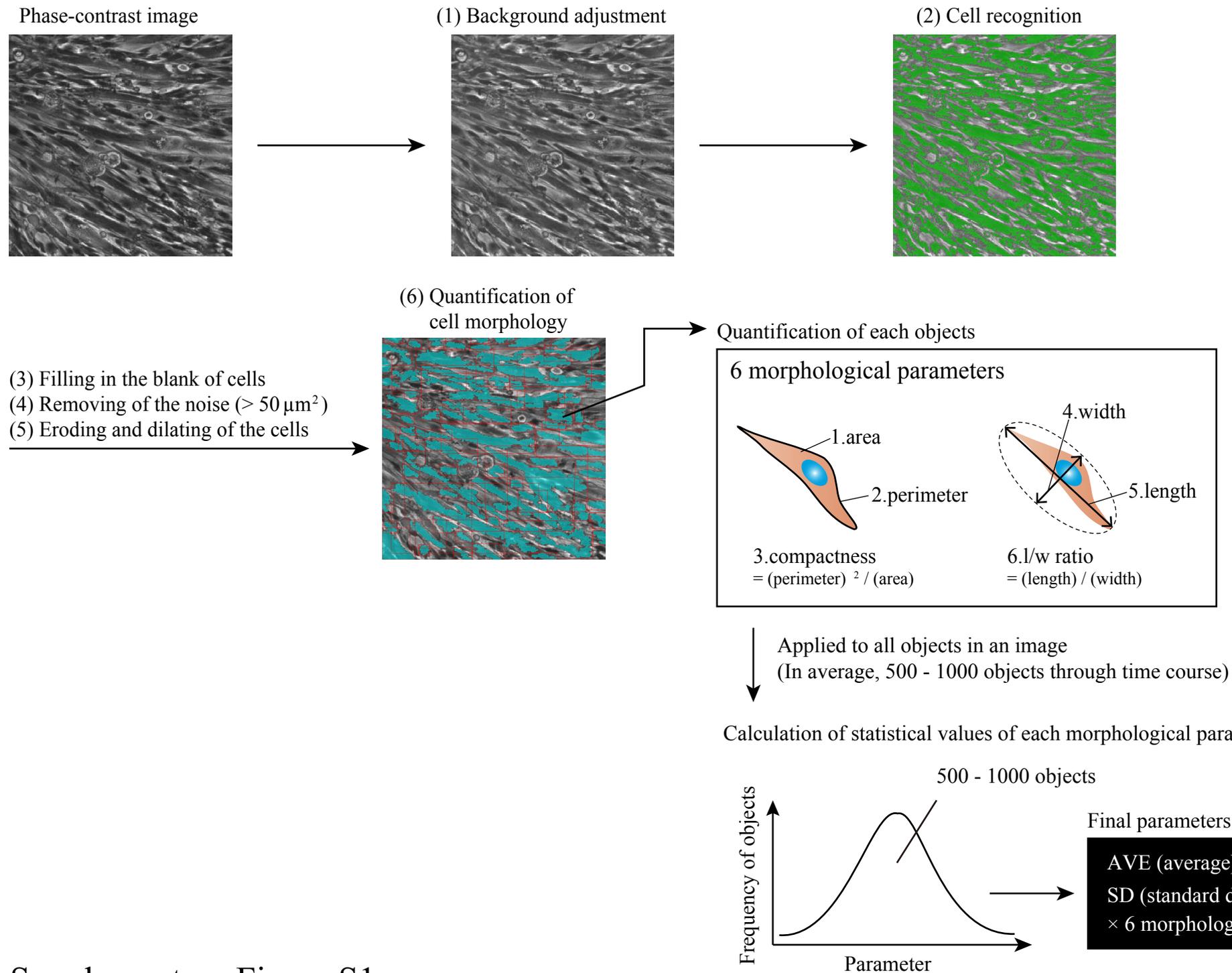
Figure S2. Prediction of an MHC-positive cellular area on days 10, 18, and 22.

(A) Representative phase contrast images and immunofluorescent images of C2C12 cells stained with MHC (green) and nuclei with SYTOX Blue (blue) on days 10, 18, and 22. The cells were

maintained under three types of culture conditions (A, B, and C). Scale bar, 400 μm . (B) Performance on the prediction of an MHC-positive area. Prediction of days 10, 18, and 22 was trained by morphological profiling from 0–10, 0–18, and 0–22 days, respectively. The dots indicate root mean square error (RMSE) throughout the analysis using different time periods of morphological information. For example, a blank dot indicates the RMSE of the prediction model trained with 0–6 hours of images. X-axis: duration of collecting the data for a morphological profile for prediction; Y-axis: RMSE values. The dotted line indicates the predictive performance of the NULL model, which is the negative control.

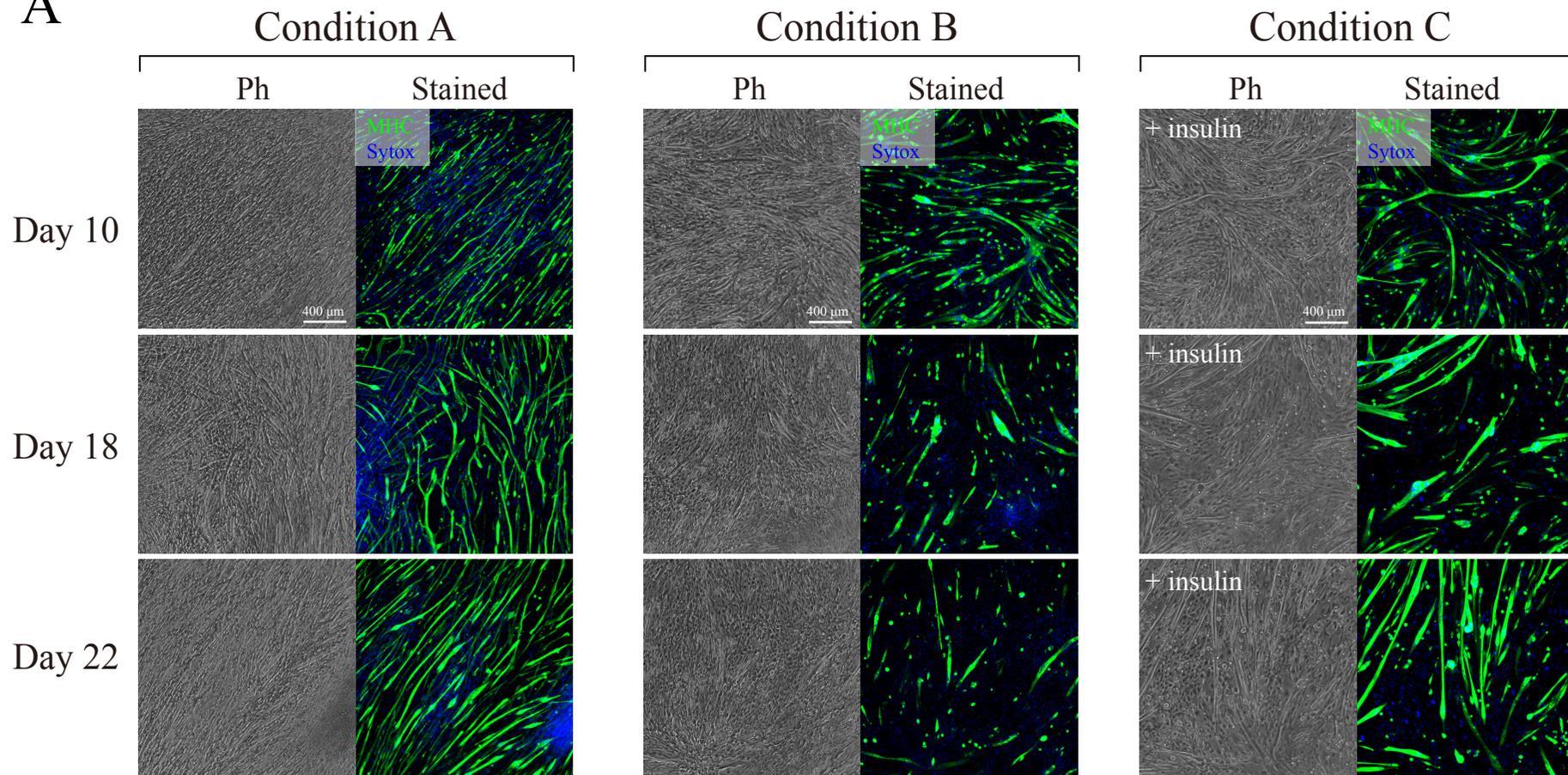
Figure S3. Analysis of a morphological response to the cGMP addition (on day 18) in condition B.

(A) Representative phase contrast images and immunofluorescent images of C2C12 cells stained for MHC (green) and nuclei by SYTOX Blue fluorescent dye staining (blue) on day 22. Scale bar, 400 μm . (B) The effect of cGMP addition on day 18 as quantified by means of the MHC-positive area. Error bars indicate means \pm SD for four independent images per C2C12 lot (total number of images = 12). (C) A morphological profile indicating the cellular response of myotubes to cGMP. The heatmap shows the AVE and SD values of the l/w ratio at each time point (6 h intervals).

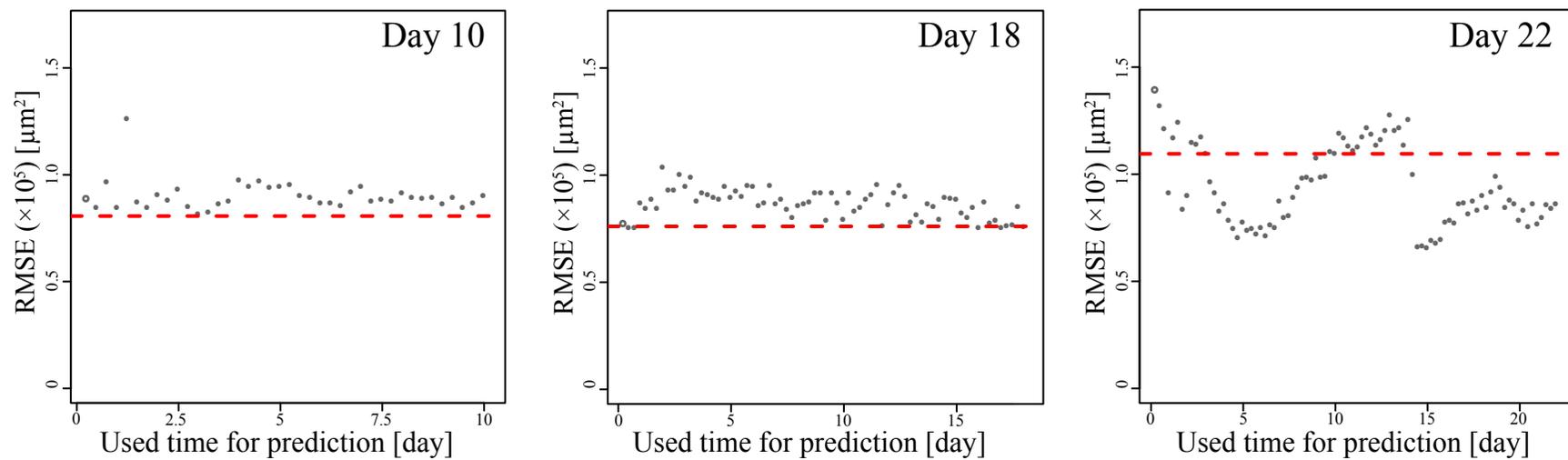


Supplementary Figure S1

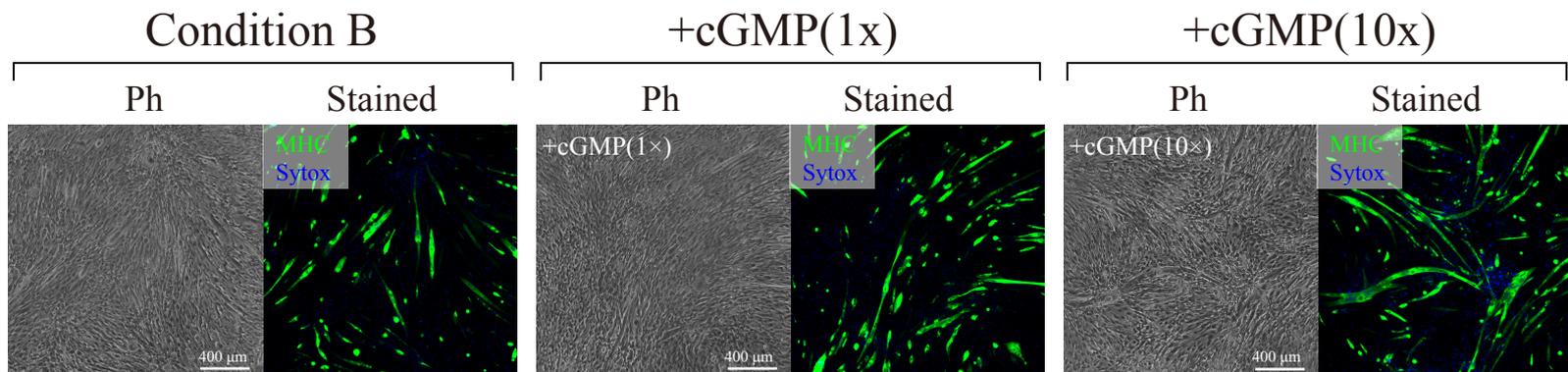
A



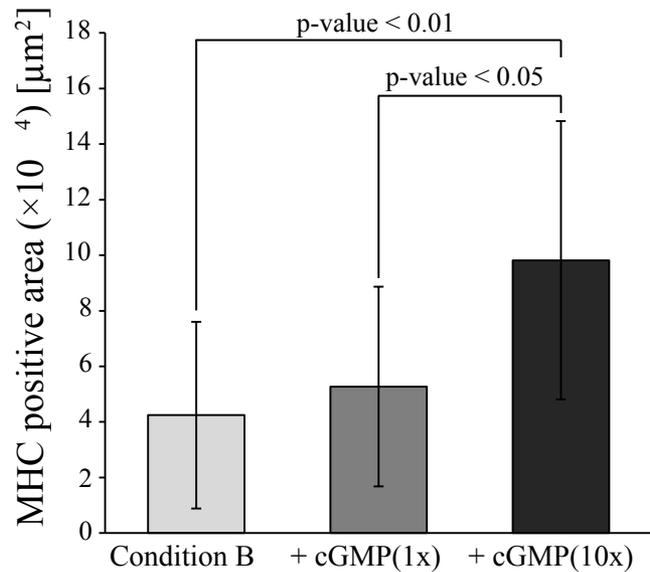
B



A



B



C

