OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Robotic Search and Early Prediction Models for Accelerating Stem Cell Manufacturing Process Development
Author(s)	都築, 拓
Citation	大阪大学, 2026, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/105597
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

THE UNIVERSITY OF OSAKA

PHD THESIS

Robotic Search and Early Prediction Models for Accelerating Stem Cell Manufacturing Process Development

---

Author :
Taku Tsuzuki

Supervisor:
Makoto Taiji

March, 2026

Abstract

Manufacturing process development in regenerative medicine is governed by the principle that **“the process is the product”**: cell-product quality cannot be fully specified by end-product testing and is therefore effectively defined by manufacturing conditions and execution history. Despite its strategic value for competitiveness, regulatory process understanding, and IP protection, systematic process development remains difficult because it faces a triple barrier: (i) high dimensionality from many interacting controllable factors, (ii) high experimental noise driven by subtle physical perturbations and operator variability, and (iii) long feedback latency in long-term differentiation protocols. These constraints often force reliance on expert intuition and trial-and-error, limiting speed, reproducibility, and scalability.

This dissertation addresses these challenges by proposing and validating a **Closed-Loop Scientific Discovery System (CL-SDS)** for regenerative medicine process development: a cyber-physical loop that integrates robotic physical execution for high-precision, massively parallel experimentation with AI-driven experimental design for sample-efficient search in high-dimensional spaces. The central stance of this thesis is that **CL-SDS is not merely an automation tool, but a necessary implementation to make tractable process development problems in which high dimensionality, high noise, and long latency coexist.**

Chapter 2 demonstrates cyber-physical loop closure in hiPSC-to-retinal pigment epithelium (RPE) differentiation. A seven-parameter search space was defined on the protocol, and a humanoid laboratory robot was integrated with batch Bayesian optimization to autonomously plan and execute experiments. Across three rounds of 48-parallel conditions, the system discovered a protocol that improved the pigmentation score—a proxy for differentiation efficiency—by 88% relative to the pre-optimized baseline, and downstream assays confirmed key marker expression and functional criteria of the resulting RPE cells.

Chapter 3 mitigates time latency by establishing early, non-destructive prediction of terminal outcomes from intermediate/non-destructive observation. In a directed hiPSC-to-muscle stem cell (MuSC) protocol with terminal quality defined as MYF5 positivity on day 82, a phase-contrast imaging–based predictor was built using images from days 14–38. Terminal outcomes were predictable more than 50 days in advance: low-efficiency outcomes (MYF5+% < 10%) were best detected at day 24 (AUC = 0.76), and high-efficiency outcomes (MYF5+% ≥ 20%) at days 31–34 (AUC = 0.86). As an operational filter, the model reduced the defective sample rate by 43.7% and increased the proportion of high-quality samples by 72%.

Chapter 4 integrates these contributions and discusses two system-level extensions enabled by combining physical loop closure with intermediate-state observability: accelerated discovery via multi-fidelity closed-loop optimization using early predictions as low-cost signals, and adaptive manufacturing via feedback control based on intermediate state estimation to suppress quality variability under patient/lot-dependent uncertainty. Overall, this work establishes CL-SDS as a practical foundation for transforming regenerative-medicine process development into a computable and controllable engineering discipline.

Table of Contents

Abstract **2**
Table of Contents **3**
Abbreviations **4**
1. Introduction **5**
1.1 Chapter Summary 5
1.2 Closed-Loop Scientific Discovery System (CL-SDS) 6
1.3 The history of Closed-Loop Scientific Discovery System (CL-SDS) 10
1.4 Process Development in Regenerative Medicine as a Frontier: Positioning, Significance and Necessity 16
1.5 Research Challenges and Thesis Outline 20
References 23
2. Robotic search for optimal cell culture in regenerative medicine **29**
2.1 Chapter Summary 29
2.2 Note on Contribution 29
2.3 Introduction 30
2.4. Results 33
2.5 Discussion 48
2.6 Materials and Methods 49
References 68
3. Early and non-destructive prediction of the differentiation efficiency of human induced pluripotent stem cells using imaging and machine learning **71**
3.1 Chapter Summary 71
3.2 Note on Contribution 71
3.3 Introduction 72
3.4 Results 75
3.5 Discussion 84
3.6 Methods 88
References 92
4. General Discussion and Conclusion **99**
4.1 Chapter Summary 99
4.2 Summary of Contributions 100
4.3 Discussion: Toward Accelerated & Adaptive CL-SDS 101
4.4 Conclusion 107
References 108
Acknowledgement **110**

Abbreviations

AI Artificial Intelligence
ARD Automatic Relevance Determination
ARD-RBF Automatic Relevance Determination Radial Basis Function (kernel)
AUC Area Under the (ROC) Curve
BCLP Batch Contextual Local Penalization
BO Bayesian Optimization
CQA(s) Critical Quality Attribute(s)
COGs Cost of Goods (Sold)
CL-SDS Closed-Loop Scientific Discovery System
CRISPR-Cas9 Clustered Regularly Interspaced Short Palindromic Repeats – Cas9
DBTL Design–Build–Test–Learn
DL Detachment pipetting length
DMD Duchenne muscular dystrophy
DOE Design of Experiments
DP Detachment trypsin period
DS Detachment pipetting strength
EI Expected Improvement
EGO Efficient Global Optimization
FCM Flow Cytometry
FFT Fast Fourier Transform
FGFRi Fibroblast Growth Factor Receptor inhibitor
GPR Gaussian Process Regression
hiPSC human induced Pluripotent Stem Cell
ICC Immunocytochemistry
iPSC induced Pluripotent Stem Cell
KP KSR concentration reducing period
KSR KnockOut Serum Replacement
MuSC(s) Muscle Stem Cell(s)
MHC Myosin Heavy Chain
MYF5 Myogenic factor 5
MYOD1 Myogenic differentiation 1
PBS Phosphate-Buffered Saline
D-PBS Dulbecco's Phosphate-Buffered Saline
PBS-T Phosphate-Buffered Saline with Tween (commonly Tween-20)
PC Preconditioning concentration
PP Preconditioning period
qRT-PCR quantitative Reverse Transcription PCR
RBF Radial Basis Function
ROC Receiver Operating Characteristic
RPE Retinal Pigment Epithelium
RSM Response Surface Methodology
RT-qPCR Reverse Transcription quantitative PCR
3P Three-chemical (Y, SB, CKI) supplement administration period
Y Y-27632 (ROCK inhibitor)
SB SB431542 (TGF-β pathway inhibitor)
CKI Casein kinase inhibitor (e.g., CKI-7)

Chapter 1

1. Introduction

1.1 Chapter Summary

Manufacturing process development in regenerative medicine represents a critical yet intractable frontier, currently bottlenecked by human limitations in physical reproducibility and cognition/planning in high-dimensional space. While the "process is the product" principle makes process optimization essential for quality and competitiveness, the triple barrier of high dimensionality, physical sensitivity, and time latency has kept this domain as an artisanal craft. Against this background, this chapter establishes the **Closed-Loop Scientific Discovery System (CL-SDS)** not merely as a conceptual framework, but as a strategic methodology to **transform "process development in regenerative medicine" from an artisanal craft into a computable engineering discipline.** This chapter is structured to logically connect the general framework to the specific contributions of this thesis:

(1) Defining the Framework (Sections 1.2 & 1.3)

I first define CL-SDS as a cyber-physical loop that integrates Robotics (for "high precision with massive parallelism") and AI (for "high-dimensional strategic planning"). By reviewing the genealogy from Design of Experiments (DOE) to modern laboratory automation, I argue that the convergence of these technologies is a historical necessity to overcome the cognitive and physical limits of manual experimentation.

(2) Positioning the Frontier (Section 1.4)

Crucially, I identify manufacturing process development in regenerative medicine as the "unexplored frontier" (white space) where CL-SDS creates the most significant value. I argue that this domain is governed by the principle that "the process is the product," yet remains intractable due to a triple barrier: vast search spaces ("High Dimensionality"), extreme sensitivity to physical handling ("Physical Sensitivity"), and prohibitive experimental costs/latency ("Time Latency") . This section frames the optimization of cell manufacturing not just as an efficiency problem, but as a fundamental scientific challenge that requires robotic autonomy to solve.

(3) Defining the Thesis Contributions (Section 1.5)

Against this frontier, I define the specific roles of the subsequent chapters as a two-pronged solution to these barriers. **Chapter 2** addresses the barriers of "High Dimensionality" and "Physical Sensitivity." I demonstrate the first physical implementation of CL-SDS in this domain, achieving highly reproducible process optimization using humanoid robots and Bayesian optimization. **Chapter 3** addresses the barrier of "Time Latency." I introduce a non-destructive predictive model using imaging and machine learning to forecast outcomes months in advance, thereby accelerating the discovery cycle.

1.2 Closed-Loop Scientific Discovery System (CL-SDS)

This section defines the **Closed-Loop Scientific Discovery System (CL-SDS)** as a cyber-physical system that autonomously iterates Model Generation, Experimental Design, and Physical Execution under a user-specified Problem Definition and Search Strategy. I first establish a precise definition and terminology in Section 1.2.1, and then argue why closing the loop is essential for modern, high-dimensional and noise-sensitive experimental problems in Section 1.2.2: robotics provides reproducible, massively parallel data generation, and AI provides high-dimensional strategic planning. The historical development of these ideas and technologies is reviewed in Section 1.3.

1.2.1 Defining the Closed-Loop Scientific Discovery System (CL-SDS)

This section formally defines the "Closed-Loop Scientific Discovery System (CL-SDS)," which serves as the core framework of this chapter (**Figure 1.1**). The attempt to redefine the scientific discovery process as a computational algorithm reproducible on a machine situates itself within the historical lineage of philosophy of science—ranging from Francis Bacon's advocacy of inductive reasoning[1] to Karl Popper's arguments on falsifiability[2]. The CL-SDS can be viewed as a physical implementation of these classical scientific methodologies, realized through modern robotics and artificial intelligence.
To ensure rigor in our discussion, this thesis defines the CL-SDS as an integrated system of computation and physical instrumentation that satisfies the following structural components and functional requirements.

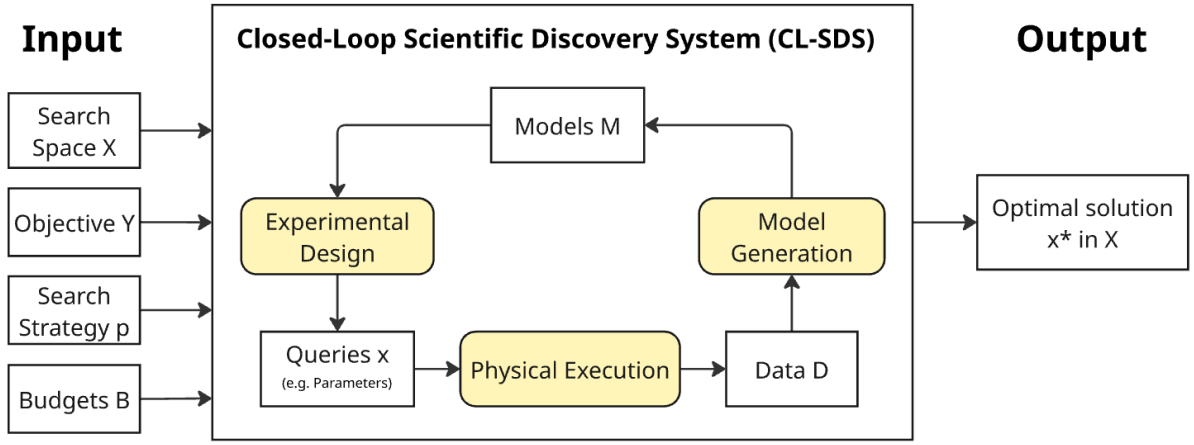


Figure 1.1: The Architecture of a Closed-Loop Scientific Discovery System (CL-SDS)
The basic architecture of a CL-SDS. The system receives a "Problem Definition" and "Search Strategy" as external inputs and returns an optimal solution x^* in search space X. The Problem Definition prescribes the "Search Space" (including Constraints C), "Objective," and "Budgets." Internally, three functional blocks—"Model Generation," "Experimental Design," and "Physical Execution"—form a closed loop. Data obtained from Physical Execution updates the Model (Induction), and based on the updated Model, a Query (next experimental condition) is generated (Deduction) and sent to Physical Execution

(1) Definition:

A Closed-Loop Scientific Discovery System (CL-SDS) is defined as a system that, when provided with an external Problem Definition and Search Strategy, autonomously iterates the internal functional cycle of (1) Model Generation (Induction), (2) Experimental Design (Deduction), and (3) Physical Execution (Observation) to acquire unknown knowledge x^* (e.g., global optima, improved parameter) within a search space without human intervention. This autonomous loop integrates "Induction," which derives laws from observations, "Deduction," which derives events to be verified from hypotheses, and physical intervention into the real world.

(2) Components:

The static information elements constituting the CL-SDS are defined as follows:

1. Problem Definition: A specification prescribing the framework for search and verification. It is the input that determines "what the system should solve" and consists of three main elements:

- **Search Space (X):** The set of executable experimental conditions. This is a multi-dimensional space defined by the feasible ranges of parameters such as medium composition, temperature, and culture duration. It also includes Constraints (C), which define exclusion zones or boundary conditions based on physical resource limits, safety restrictions, or biological viability, thereby determining the effective Feasible Region.
- **Objective (y):** The goal, evaluation criterion, or reward function of the experiment. It is a metric that interprets data obtained from experiments to quantify whether an experiment was a "success" or how valuable it was (e.g., differentiation efficiency, pigment production, purity).
- **(Optional) Budgets (B):** Constraints prescribing the upper limits of resources available for the search. These include the total number of experiments (Total Iterations), the number of experiments that can be executed in parallel (Batch Size), time limits, or financial budgets. These dictate to the Search Strategy when to terminate the search or what set of experiments to propose at one time.

2. Search Strategy (P): An algorithm or policy for selecting the next experimental point within the search space X to maximize (or minimize) the objective y. It prescribes what model and what method (e.g., acquisition function) are used to generate the plan.

3. Model (M): A mathematical representation (hypothesis) describing the relationship between input X and output y, constructed based on past experimental data. This ranges from simple linear regression to complex non-linear models like Gaussian process regression or neural networks. The system may maintain multiple models.

4. Query (x_{next}): Specific instructions for the next experimental condition to be executed. Output by the Experimental Design function, this serves as the input to the Physical Execution function. It takes a form convertible into specific robotic commands.

5. Data (D): A set of observations obtained as a result of physical experiments. Accumulated as pairs of input conditions x and corresponding responses.

(3) Functions:

The dynamic processes executed by the CL-SDS to drive the cycle are defined as follows:

Model Generation: The process of extracting underlying regularities or causal structures from accumulated data D to construct or update the model M. This assumes the function of "induction," estimating general laws (or local response surfaces) from individual observational facts.

Experimental Design: The process of determining "what to observe next" to achieve the objective y, based on the current model M and search strategy P. This corresponds to the process of "deduction," deriving experimental conditions to verify hypotheses. Computationally, this is formulated as an optimization problem, such as maximizing an acquisition function, to output the next query x_{next}.

Physical Execution: The process of receiving the query x_{next}, performing actual experimental operations in the Physical World, and acquiring new data D. This includes pipetting by robotic arms, cell culture, and sensing. Unlike simulations (in silico) where everything concludes within the computer, this process is characterized by the accompaniment of real-world noise, uncertainty, and physical costs.

1.2.2 Motivation: Why “Robot-AI Closed-Loop” ?

If the definition of CL-SDS is the integration of "Model Generation," "Experimental Design," and "Physical Execution," a fundamental question arises: **Why is this integration necessary?** Is it merely to accelerate the speed of experiments, or is there a more essential scientific motivation? This section argues that the necessity of the Closed-Loop system stems from two transformative capabilities that revolutionize modern scientific discovery: "High Precision with Massive Parallelism" provided by Robotics and "High-Dimensional Strategic Planning" provided by AI.

(1) Robotics: Massive Parallelism in High Precision

The first motivation lies in the physical "body" of science. While the "Reproducibility Crisis" [3] highlights the inherent variability of manual experimentation, the defining strength of Robotics in a CL-SDS is its ability to execute **“High-Precision” experiments with "Massive Parallelism."** In manual experimentation, maintaining consistency across hundreds of samples is physically impossible due to human-induced variability such as subtle differences in pipetting speed or agitation. This variability is not trivial; Kanie et al. (2019) demonstrated that parameters related to physical manipulation can critically alter experimental outcomes[4]. Robotics overcomes this by enhancing "Reproducibility" [5], ensuring that every physical operation is executed with identical parameters by reducing human-induced variability. Crucially, robotics enables "Scale" without compromising this precision: it can run tens-to-hundreds of conditions simultaneously in parallel. This capacity to consistently generate low-noise data at a massive scale provides the reliable ground truth necessary for AI algorithms to work effectively.

(2) AI: High-Dimensional Cognition and Strategic Planning

The second motivation lies in the computational "brain" of science. As scientific questions deepen, the Search Space X becomes inherently high-dimensional due to the simultaneous interaction of numerous parameters. This combinatorial explosion creates a critical "cognitive bottleneck" [6]: while humans can intuitively grasp correlations between limited variables, it is cognitively challenging to navigate the complex, multi-dimensional landscape of the objective function $f : \mathcal{X} \to y$ without computational aid, often reducing manual discovery to a matter of serendipity.More critically, humans lack the processing capability to coherently design massive parallel experiments. While a researcher might intuitively plan a single experiment based on a hypothesis, generating a batch of dozens or hundreds of distinct experimental queries that are strategically balanced is beyond human cognitive capacity. Artificial Intelligence (AI), such as batch Bayesian optimization [7], provides this strategic planning capability. AI can not only model high-dimensional interactions that human intuition misses but also generate optimal, non-redundant experimental plans for massive parallel execution, maximizing the efficiency of the robotic search.

(3) The Synergy: Efficient Exploration of the Vast/High-Dimentional Search Space

The essence of the CL-SDS lies in the synergy of combining these two solutions. AI algorithms, no matter how sophisticated, cannot learn the correct laws if the input Data D contains excessive noise. If the input data is compromised by variability, the AI struggles to distinguish true signal from noise. This degradation validates the adage "Garbage In, Garbage Out," potentially negating the strategic advantage of AI and rendering the search process comparable to blind trial-and-error.

The "Closed-Loop" structure overcomes this limitation by coupling **"High Precision with Massive Parallelism"** provided by Robotics with **"High-Dimensional Strategic Planning"** provided by AI.

- **Robots provide "High Precision with Massive Parallelism"** ensuring high-fidelity data generation at a scale challenging for humans, thereby lowering the noise floor for AI.
- **AI provides "High-Dimensional Strategic Planning,"** efficiently navigating complex search spaces beyond human cognitive limits to allocate the robot's limited query budget to the most promising and informative conditions.

This study posits that for scientific problems characterized by high-dimensional search spaces, extreme sensitivity to physical variability, and high per-query costs that impose strict limits on the total number of feasible experiments, the CL-SDS provides value far beyond simple experimental efficiency. In these domains, this synergy is not merely an accelerator but a scientific necessity—an indispensable approach to uncover knowledge that lies beyond the reach of human intuition and manual execution. As elaborated in Section 1.4, process development in regenerative medicine is an archetype of such a domain. The manufacturing process can involve optimizing dozens or even more interacting parameters, a complexity that exceeds human cognitive limits. Moreover, the biological outcomes are highly sensitive to manual handling variability, and with single experiments often requiring over a month to complete—thereby constraining the total query budget—rigorous process development is difficult without a system capable of executing massive parallel exploration with absolute physical consistency. To demonstrate this capability, Chapter 2 presents a

proof-of-concept study applying CL-SDS to the differentiation of retinal pigment epithelial (RPE) cells from human induced pluripotent stem cells (hiPSCs). The results establish that CL-SDS can autonomously navigate the complex search space to discover novel protocols yielding differentiation efficiencies.

1.3 The history of Closed-Loop Scientific Discovery System (CL-SDS)

This section reviews the historical genealogy of the CL-SDS defined in Section 1.2. Rather than viewing automation as a monolithic trend, I analyze history as the independent maturation of its key functions and their eventual integration. Section 1.3.1 traces the evolution of these components, ranging from the statistical foundations of Design of Experiments (DOE) to the rise of laboratory robotics, culminating in the first physical implementation of the CL-SDS by the Robot Scientist "Adam"[8]. Subsequently, Section 1.3.2 surveys the rapid proliferation and diversification of CL-SDS across materials science and life sciences from 2010 onwards.

1.3.1 The Road to the First Closed-Loop Scientific Discovery System

As defined in Section 1.2.1, a Closed-Loop Scientific Discovery System (CL-SDS) is a system that, given externally specified Problem Definition (including at least Search Space X, Objective y, and Budgets B and a Search Strategy P, can autonomously iterate the internal cycle of (1) Experimental Design → (2) Physical Execution → (3) Model Generation, thereby producing new knowledge x^* in the search space X. This subsection places a historical time axis onto that definition: rather than recounting “automation” as a single linear trend, I track which CL-SDS functions were automated first, which remained human-dependent, and what finally enabled physical loop closure in the 21st century.

(1) Statistical Design of Experiments (DOE): The Conceptual Loop

The conceptual foundation of the discovery loop predates modern AI and can be traced to the formalization of Statistical Design of Experiments (DOE). Established by R. A. Fisher in 1935, DOE provided a rigorous mathematical framework for planning experiments (Experimental Design) that support valid inference within a noisy search space X under a predefined objective y [9]. This framework was further extended by Response Surface Methodology (RSM) published in 1951, which explicitly framed the optimization of unknown processes as an iterative, model-assisted search over X [10]. These traditions demonstrate that the core logic of a Closed-Loop Scientific Discovery System was conceptually present by the mid-20th century. However, when interpreted through the formal CL-SDS definition introduced in Section 1.2.1, DOE-based workflows remained fundamentally Human-in-the-loop systems. While Experimental Design was mathematically guided and the search space X and objective y were explicitly defined, the rotation of the loop itself was not autonomous. Physical Execution relied entirely on manual labor, and Model Generation—typically fitting low-order response surfaces—required human judgment to interpret results and determine subsequent experimental directions. Thus, the missing step for autonomy was not the idea of sequential experimentation, but the automation and integration of Physical Execution and Model Generation required to repeatedly rotate this loop without human intervention.

(2) Inductive Machines: Automated Model Generation that did not act

With digital computing, one of the first widely recognized attempts to mechanize scientific reasoning focused on Model Generation from Data D. Representative systems include:

- **DENDRAL(1978):** demonstrated that computers could infer molecular structure hypotheses from mass spectrometry data in specific domains [11].
- **BACON(1977):** rediscovered empirical laws from observational data, showing that rule-like structure could be induced from measurements [12].
- **Symbolic regression for natural laws(2009):** learned compact analytic forms of dynamical laws from data, reinforcing that “model discovery” could be automated beyond narrow expert systems [13].

These systems performed nontrivial inference and hypothesis construction. The sharper limitation, in CL-SDS language, is that they generally assumed Data D arrived as an external input. They did not, as a system, autonomously run Experimental Design to propose a next query x, nor did they integrate Physical Execution to generate new data. Put differently: they mechanized **Model Generation**, but not **Model → Experimental Design → Physical Execution → Data** as a single autonomous loop.

(3) Discovery as a Search: from heuristic problem spaces to sequential optimization

In parallel with the development of inductive machines, AI and decision science established the theoretical framework necessary to define the Problem Definition, Search Strategy P and Experimental Design in a CL-SDS. A pivotal moment was Newell and Simon’s conceptualization of intelligent problem-solving [14]. They framed intelligence not as vague intuition, but as the heuristic navigation of a "hypothesis space" to reach a goal state. In the context of the CL-SDS, this perspective fundamentally redefined scientific discovery: it is no longer an abstract intellectual activity, but a concrete search problem to find an optimal solution x^* within a defined Search Space X. Consequently, Experimental Design became formalized as a sequential decision-making process: applying a Search Strategy P to select the most promising next query x based on the current Model M.

This “discovery as a search” view matured first in computer-aided design/engineering/manufacturing (CAD/CAE/CAM), where the loop between Model Generation and Experimental Design could be closed entirely in silico by using simulations as the evaluation oracle. During the 1990s, optimization algorithms such as evolutionary computation became practical engines for such digital closed loops, particularly in domains where evaluations were computationally feasible or parallelizable: for example, hybrid evolutionary optimization coupled with CFD for airfoil and wing design [15], genetic-programming-based synthesis of analog electrical circuits satisfying complex performance specifications [16], and genetic-algorithm applications to molecular recognition problems including ligand docking [17]. These systems were “closed loop” in the computational sense, but they were not physical CL-SDSs: the Physical Execution step was replaced by simulation, largely insulated from the noise, drift, and latency that dominate wet-lab experimentation. For a physical CL-SDS, the decisive bottleneck is sample efficiency under tight experimental budgets B: how to choose the next physical trial x when each evaluation

is expensive, noisy, and time-consuming. This regime favors surrogate-based sequential optimization, in which Model Generation builds a probabilistic surrogate from accumulated Data D, and Experimental Design uses that surrogate to propose the next query by maximizing an acquisition criterion. Bayesian optimization—introduced as a Bayesian approach to global optimization aimed at minimizing the number of costly function evaluations [18] and later formalized for engineering design as Efficient Global Optimization (EGO) using Kriging/Gaussian-process surrogates and acquisition functions [19]—thus provides the mathematical engine for the Experimental Design ↔ Model Generation loop. Although originally designed for deterministic functions, its Gaussian process framework naturally extends to handle the noise inherent in physical experimentation. This is the core algorithmic idea underpinning the robotic search system in regenerative medicine developed in Chapter 2 of this thesis.

Clarification on representation: In Newell and Simon's framework, the "search space" (problem space / hypothesis space) is an abstract construct whose concrete form depends on how candidate solutions are represented. In this thesis, I focus on a practically important instantiation in experimental science: representing a candidate solution as an executable experimental condition parameterized by a finite set of controllable variables. Under this representation, **"discovery as a search" becomes optimization in a design/parameter space.** Optimization is thus a special case of search in which candidate solutions are represented by a parameter vector and evaluated by a scalar objective. Depending on whether the controllable variables are discrete (e.g., reagent choices, timing options) or continuous (e.g., concentrations, durations), this instantiation corresponds to combinatorial search or continuous numerical optimization. In both cases, the core difficulty is sample-efficient sequential experimentation under tight budgets.

(4) Laboratory Automation: "Open-loop" automation of Physical Execution

Parallel to the computational advances in "search," the physical "body" of science evolved from manual craft into a system of programmable Physical Execution. This transformation occurred through two pivotal phases: the algorithmic standardization of individual operations and the subsequent scaling of generalized robotics. Early breakthroughs demonstrated that wet-lab processes could be reduced to machine-executable instructions. Skeggs (1957) standardized measurement via continuous-flow analysis, proving that assays could be executed as a continuous stream rather than discrete handling[20]. Merrifield (1963) revolutionized synthesis with Solid-Phase Peptide Synthesis (SPPS), which redefined complex chemical production as an iterative, programmable cycle of "Add–React–Wash" [21]. These innovations established the conceptual template for treating physical experiments as code.

During the 1980s, automation expanded beyond dedicated instruments toward more general-purpose laboratory robotics. A commercially visible milestone was the Zymark Zymate laboratory automation system, which exemplified a reprogrammable robotic platform capable of executing diverse workflows once operations were specified [22]. As standardized microplates, liquid-handling devices, and plate readers matured, these components were integrated into programmable "workcells," enabling large-scale execution of parameterized assays. This technological base underpinned the rise of high-throughput screening (HTS) in pharmaceutical R&D: by the late 1980s and early 1990s, HTS-style workflows had operational roots in microplate-based, miniaturized assays and were rapidly scaled as

robotics and instrumentation matured[23,24] . However, from the perspective of CL-SDS, canonical HTS pipelines largely remained open-loop: experimental protocols were pre-specified by humans, and subsequent Queries were not adaptively updated through Model Generation and Experimental Design under an explicit Search Strategy. In short, HTS delivered an exceptionally powerful “body” for Physical Execution and rapid data acquisition, but not yet the integrated closed-loop architecture required for autonomous discovery.

(5) The Robot Scientist “Adam”: The integration point as a CL-SDS milestone

The development of the Robot Scientist “Adam” by King et al. (2009) is widely cited as an early wet-lab prototype in which the Closed-Loop Scientific Discovery System (CL-SDS) architecture became a physical reality[8]. Adam integrated computational decision-making (the “brain”) with laboratory robotics (the “body”) to implement an autonomous cycle of Model Generation, Experimental Design, and Physical Execution under a concrete Problem Definition.

- **Model Generation:** Using a logical model of yeast metabolism together with bioinformatic resources, the system abductively generated hypotheses linking candidate genes to orphan enzyme activities and updated their evidential status from experimental Data D.
- **Experimental Design:** It deduced discriminative microbial growth experiments selecting deletion strains and defined media with controlled metabolite supplementation in order to separate competing hypotheses.
- **Physical Execution:** Laboratory robotics executed the microplate batch growth assays, collected phenotype time series data, and automatically fed the results back to refine the model.

Crucially, Adam demonstrated that this cyber-physical closed loop could function without human intervention. In a case study identifying genes encoding specific orphan enzymes, the system autonomously formulated hypotheses, executed experiments to test them, and successfully identified the correct gene functions. The authors subsequently confirmed these autonomous discoveries via manual assays, proving that a CL-SDS could generate and verify scientific knowledge comparable to human researchers.

1.3.2 The Cambrian Explosion of CL-SDS (2010~)

In 2009, “Adam,” developed by Ross King et al., achieved a monumental milestone in the history of science as a system that autonomously completed the scientific discovery cycle—from Model Generation to Experimental Design, Physical Execution, and interpretation of results—within a formally specified domain [8]. From the 2010s to the present, starting with Adam, the paradigm of the Closed-Loop Scientific Discovery System (CL-SDS) has witnessed explosive diversification and diffusion across disciplines, driven by improvements in computational power, the commoditization and enhanced capabilities of robotics hardware, and the maturation and democratization of machine learning algorithms such as Bayesian optimization. This section surveys the genealogy of this evolution.

1.3.2.1 General Experimental Sciences: Materials Science and Synthetic Chemistry

Materials Science was one of the earliest and most successful domains for CL-SDS adoption. In the exploration of new materials, the Search Space—consisting of relationships among Composition, Process, Structure, and Property—is vast due to its combinatorial diversity. Approaches equivalent to CL-SDS developed rapidly to navigate this nearly infinite Search Space. The Autonomous Research System (ARES) (2016), developed by Benji Maruyama et al., integrated real-time in situ Raman spectroscopy with a learning-driven planner. It autonomously identified optimal carbon nanotube growth conditions in just hundreds of experiments, accelerating a process that previously took human researchers months [25]. The Mobile Robotic Chemist (2020), developed by Andrew Cooper et al., demonstrated the potential for autonomy without infrastructure modification by using an autonomous mobile robot to operate standard laboratory equipment in a human environment. In a photocatalyst search, this system operated continuously for 8 days, performing 688 experiments and autonomously discovering a catalyst formulation with approximately six times the activity of the baseline [26]. Project Ada (2020) optimized additive formulations (e.g., dopant concentration) for spiro-OMeTAD, a hole transport material used in perovskite solar cells. Using Bayesian optimization, it identified high-performance conditions in far fewer experiments than traditional grid searches, while also simultaneously optimizing process conditions such as annealing parameters [27].

Parallel to materials science, the application of CL-SDS-equivalent methods expanded in **Synthetic Chemistry** dealing with small molecules. The Chemputer (2019), developed by Lee Cronin et al., embodied the concept of "Chemputation," treating chemical synthesis as a universal computational process. By describing chemical reaction operations in a hardware-independent language (XDL) and executing standardized unit operations, the system succeeded in the autonomous synthesis of three pharmaceutical compounds, demonstrating that synthesis procedures can be stored, shared, and executed as digital data like software [28,29]. The MIT Platform (2019), built by Connor Coley et al., is an example integrating Design by AI and Make by robots. In this system, AI proposes synthesis routes based on retrosynthetic analysis, which are physically executed by flow reactors and liquid handling robots. In addition to successfully synthesizing 15 drug targets, it demonstrated a design-make-test loop to improve route synthesizability [30]. Furthermore, Clio (2022), developed by Venkatasubramanian Viswanathan et al., is an autonomous experimental system specialized in the search for non-aqueous electrolytes for Li-ion batteries. Coupled with the Bayesian optimization planner, Clio autonomously explored a vast compositional space of solvent types and salt concentrations, efficiently identifying high-conductivity electrolyte formulations [31].

1.3.2.2 Life Sciences: Advancements in the Drug Discovery Screening

In life sciences, particularly drug discovery screening, was among the earliest to witness the significant implementation of Closed-Loop Scientific Discovery Systems (CL-SDS). This early adoption is largely attributable to the historical context discussed in Section 1.3.1: programmable experimental execution systems via High-Throughput Screening (HTS) were already established by the 1990s [23,24], providing the necessary physical "body" for autonomous experimentation.

A pioneer in this field is the Robot Scientist "Eve" by Ross King (2015). Integrating physical screening with Quantitative Structure-Activity Relationship (QSAR) modeling, Eve autonomously hypothesized and verified that an existing anti-cancer drug could inhibit a specific parasite enzyme, demonstrating the economic viability and speed of AI-driven discovery[32]. Subsequently, the focus shifted from standalone screening to the autonomization of the complete Design-Make-Test-Analyze (DMTA) cycle, incorporating the synthesis "Make" phase. Platforms such as CyclOps (2019) and the recent MicroCycle (2024) have emerged, enabling rapid multi-parameter optimization by closing the loop between automated synthesis and evaluation[3334]. This trend has extended beyond small molecules to complex antibody therapeutics (biologics). In 2025, Grace et al. developed the EVA™ platform to engineer multispecific antibodies. By integrating high-throughput functional assays with Bayesian optimization, the system identified candidates outperforming clinical benchmarks with 14-fold higher efficiency than random search[35]. In parallel, the integration of advanced AI into the CL-SDS framework as Models and Search Strategies has significantly enhanced the quality of scientific exploration. In the realm of molecular design, Zhavoronkov et al. utilized deep generative models and reinforcement learning to discover novel DDR1 kinase inhibitors within just 21 days [36]. Complementing this in the evaluation phase, DeMeo et al. revolutionized phenotypic screening by integrating transcriptomics with closed-loop learning. Their approach identifies compounds that induce specific disease-relevant phenotypic changes—rather than simple target binding—improving success rates by orders of magnitude [37]. The fusion of these AI capabilities with automated synthesis robotics [38] has transformed CL-SDS into an intelligent engine capable of navigating vast chemical and biological search spaces with unprecedented efficiency.

1.3.2.3 Life Sciences: DBTL(Design-Build-Test-Learn) Cycle in Synthetic Biology

Synthetic biology has long pursued the engineering of microorganisms as “cell factories,” where genetic circuits and metabolic pathways are rationally designed to maximize target phenotypes (e.g. yield, quality). A canonical operational paradigm in this domain is the Design-Build-Test-Learn (DBTL) cycle: an explicit definition of the design space and objectives (Design), automated DNA construction and strain generation (Build), high-throughput phenotyping and measurement (Test), and model update to propose the next experiments (Learn). In system-engineering terms, DBTL constitutes a closed-loop optimization workflow, and biofoundries represent its infrastructural embodiment at scale [39,40]. A representative academic demonstration of an end-to-end DBTL closed loop is the “BioAutomata” platform reported by HamediRad et al. in 2019 [41]. Using the Illinois Biological Foundry for Advanced Biomanufacturing (iBioFAB) as the physical execution layer, BioAutomata coupled automated plasmid construction, culturing, and measurement with closed-loop experiment selection via Bayesian optimization to optimize the lycopene biosynthetic pathway. Notably, the study reports that the platform evaluated less than 1% of the possible design variants and outperformed random screening by 77%. In parallel, the 2010s saw the consolidation of DBTL at industrial scale via biofoundries and cloud laboratories. Biofoundries integrate standardized pipelines spanning design tools, DNA synthesis/assembly, transformation, colony management, phenotyping, and data infrastructure, enabling massively parallel strain engineering beyond manual artisanal workflows [39,40,42]. Moreover, “cloud labs” externalize Build/Test as a service: platforms such as Transcriptic/Strateos allow experiments to be specified as executable protocols and

carried out remotely on automated infrastructure, lowering the barrier to running DBTL cycles without owning a physical laboratory [43,44]. Emerald Cloud Lab and related providers further illustrate this decoupling of design/learning from physical execution as a scalable pathway for closed-loop experimentation [45]. Taken together, synthetic biology has arguably been one of the earliest life-science domains to institutionalize CL-SDS through DBTL.

1.4 Process Development in Regenerative Medicine as a Frontier: Positioning, Significance and Necessity

The preceding sections have argued that the Closed-Loop Scientific Discovery System (CL-SDS) is a powerful paradigm for scientific discovery and have outlined its historical genealogy. This section discusses the academic and industrial significance of applying this system to the domain of manufacturing process development in regenerative medicine. First, in Section 1.4.1, I organize the current state of technology using a matrix of "Modality" and "Value Chain," clarifying that process development in regenerative medicine is an unexplored "frontier" for CL-SDS. Next, Section 1.4.2 discusses why knowledge discovery in this domain holds high value, approaching the issue from multiple perspectives based on the unique characteristic of biologics: "The process is the product." Finally, Section 1.4.3 concludes why a CL-SDS integrating robotics and AI is not merely an efficiency tool but a necessity for solving the challenges inherent in regenerative medicine manufacturing, such as high dimensionality, high cost, and dependence on manual skills.

1.4.1 Positioning: Where is the Frontier? (The Modality × Value Chain Map)

By organizing the application areas of Closed Loop Scientific Discovery System (CL-SDS) along two axes—"Target Modality" and "Value Chain Stage"—manufacturing process development in regenerative medicine emerges as a distinct "white space" (**Figure 1.2**).

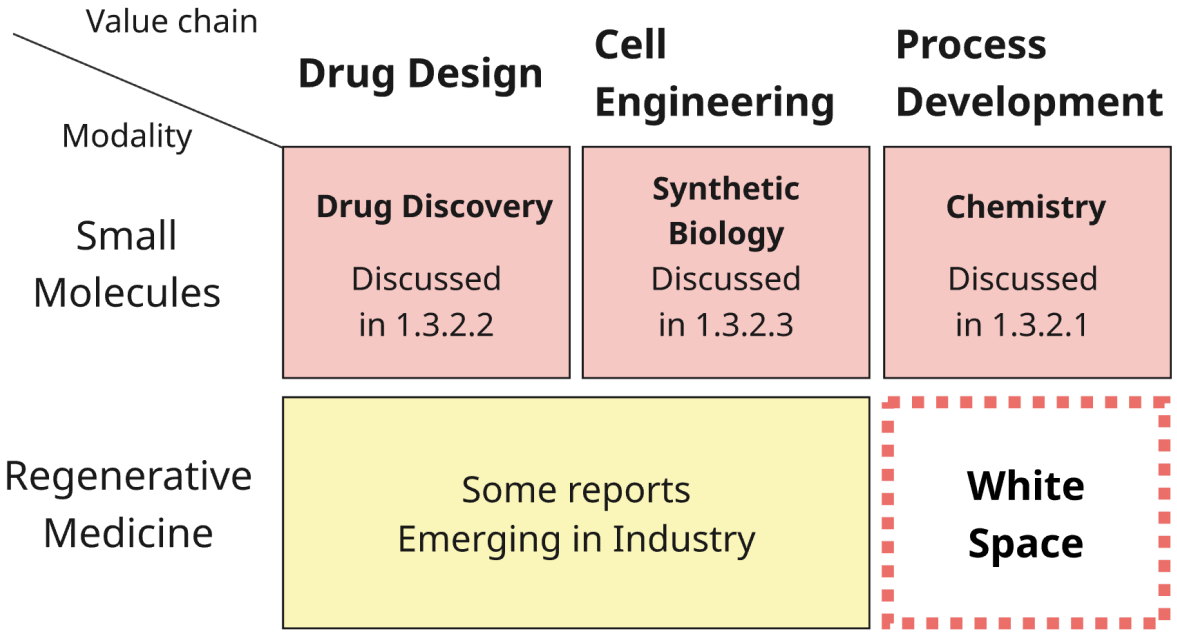


Figure 1.2: The Frontier of Closed-Loop Scientific Discovery Systems (CL-SDS).
A "Modality × Value Chain" map illustrating the maturity of CL-SDS adoption across different domains. The vertical axis represents the therapeutic modality (Small Molecules, Regenerative Medicine), and the horizontal axis represents the stage in the pharmaceutical value chain (Drug Design, Cell Engineering, Process Development). While CL-SDS is established in small molecule drug design[32] and synthetic biology[39,41], and increasingly in chemical process optimization[26,28], the application of autonomous closed-loop systems to the manufacturing process development of regenerative medicine remains a largely unexplored "white space" or frontier (red). This thesis targets this specific frontier.

---

(1) Small molecules × Drug Design: Established

As discussed in Section 1.3.2.2, the effectiveness of CL-SDS in small molecule discovery is well-established. Following the foundational work of the Robot Scientist "Adam"[8] in functional genomics, the Robot Scientist "Eve" successfully applied the CL-SDS paradigm to high-throughput screening (HTS) for drug discovery[32]. The field has since evolved into fully integrated Design-Make-Test-Analyze (DMTA) cycles, effectively closing the loop for targets with clearly defined structures.

(2) Small molecules × Cell Engineering: Established

As discussed in Section 1.3.2.3, autonomous optimization in microbial cell engineering is systematized via the DBTL (Design-Build-Test-Learn) cycle in Synthetic Biology. Industrial-scale biofoundries and cloud laboratories now serve as the physical infrastructure for this loop, enabling the rapid and systematic optimization of gene circuits and metabolic pathways [39,41].

(3) Small molecules × Process Development: Established

As discussed in Section 1.3.2.1, CL-SDS applications in process development for chemistry and materials science are advancing rapidly. Innovative platforms like the "Mobile Robotic Chemist" for organic synthesis demonstrate that "autonomy stacks" can successfully navigate the search space of synthetic pathways and reaction conditions without human intervention[26].

(4) Regenerative medicine × Drug Design / Cell Engineering: Emerging Signals

In the upstream phases of regenerative medicine—particularly engineered-cell design and functional screening—industry-led initiatives are beginning to resemble DBTL/CL-SDS-like workflows that couple computational design with high-throughput experimentation. For example, Hitachi has reported a “DesignCell” platform that integrates generative design of CAR gene sequences with high-throughput functional evaluation, combining pooled single-cell screening with a robotics-based array screening workflow that automates steps from gene introduction to cell-function analysis. These efforts constitute credible “emerging signals” of partial loop closure in engineered-cell discovery.[46]

(5) Regenerative medicine × Process Development: The Frontier

In contrast, "Process Development" in regenerative medicine remains a vast white space. While numerous studies have mechanized specific physical tasks in cell culture [47–55],

most remain at the level of "physical automation," simply reproducing predetermined procedures. Implementation as a CL-SDS that autonomously explores and improves the process conditions themselves is extremely rare. Therefore, this thesis targets this unexplored frontier. By deploying an autonomous closed-loop system integrating robotics and AI, I aim to transform this domain—traditionally dependent on manual heuristics—into a tractable scientific subject.

1.4.2 Significance: Why Manufacturing Knowledge Discovery is High-Value in Regenerative Medicine?

Why is knowledge discovery (exploration of optimal conditions) in the manufacturing process development of regenerative medicine important? The answer lies in the fundamental distinction between conventional small-molecule drugs and living cell products: the methodology for establishing product identity and quality. For small molecules, the active ingredient is a distinct chemical entity defined by its structure. Accordingly, a large portion of product identity and quality can be verified through established analytical characterization and specification-based testing, although robust process control remains essential throughout manufacturing.

In contrast, living cells are exceedingly complex systems that cannot be exhaustively characterized by current analytical technology alone—particularly for functional attributes such as potency—making the "history" of the cell just as critical as its current state. This constraint is often summarized as the principle that **"The process is the product"**: since complex cellular products cannot be exhaustively characterized by end-product testing, their identity and quality are inherently defined by the consistency of the manufacturing process itself [56–58]. Consequently, regenerative medicine strongly benefits from a prospective "**Quality by Design" (QbD)** approach, in which the causal relationships between process parameters and product quality—the "Design Space"—are explicitly established to assure safety and efficacy. In this context, knowledge discovery in process development holds exceptional value from four perspectives:

(1) Quality Significance

Variations in the manufacturing process directly dictate the efficacy and safety of the cell product. Slight deviations in culture conditions can critically alter differentiation efficiency and functionality, effectively creating a "different drug." Elucidating the causal relationship between process parameters and critical quality attributes (CQAs) is, therefore, synonymous with designing the product itself.

(2) Economic Significance

In regenerative medicine, particularly autologous therapies like CAR-T, manufacturing costs constitute a significant driver of the Cost of Goods Sold (COGs). Complex logistics, labor-intensive manual procedures, and extended culture durations drive these costs [59]. Consequently, process optimization is the primary determinant of therapeutic feasibility and scalability.

(3) Regulatory Significance

From a regulatory science perspective, a scientific understanding of the process is mandatory. The ICH Q8 guideline defines "Design Space" as the multidimensional

combination of input variables that assures quality [60]. Furthermore, assessing comparability during manufacturing changes requires a causal understanding of the process-quality relationship [58,61].

(4) Strategic/IP Significance

Given the difficulty of complete characterization, process know-how serves as a robust source of competitive advantage. Beyond patent protection, many entities protect manufacturing know-how as trade secrets, constructing competitive moats that are difficult to replicate [62].

Thus, exploring the manufacturing process in regenerative medicine is not merely procedural but a "scientific challenge worth solving" that generates substantial medical, economic, and strategic value.

1.4.3 Necessity: Why Robots-AI Closed-Loop Matters?

Why is a Robot-AI CL-SDS necessary to address this frontier? The necessity arises because the challenges inherent to this domain exceed the limits of human cognitive and physical capabilities. This section reaffirms the utility of CL-SDS discussed in Section 1.2 within the specific context of regenerative medicine.

(1) High Dimensionality（Vast Search Space)

Cell culture optimization involves a combinatorial explosion of parameters—medium composition, factor concentrations, timing, and scaffold materials. The simultaneous optimization of these interacting variables is a high-dimensional problem that is difficult to approach efficiently with manual intuition or one-factor-at-a-time experimentation; AI-driven experimental design can improve sample efficiency by prioritizing informative experiments.

(2) High Latency(Time Consuming)

Unlike high-throughput screening, cell differentiation experiments are highly time-consuming (have high latency), often requiring months per trial. Since the total number of feasible experiments is severely limited, inefficient strategies like random or grid searches are impractical. Therefore, the adoption of an efficient search strategy is indispensable.

(3) Sensitivity to Physical Manipulation

Cells are exquisitely sensitive to physical perturbations. "Variations in technique" not captured in protocols—such as pipetting mechanics, vibration, and handling speed—can critically impact differentiation outcomes [4]. In practice, maintaining perfectly consistent physical execution across long culture periods and large numbers of samples is extremely difficult for human operators. Such uncontrolled variability inflates experimental noise, thereby obscuring causal relationships between process parameters and biological outcomes. Robotic physical execution substantially mitigates this problem by standardizing low-level operations with high precision and reproducibility, reducing noise at the data-generation stage. This reduction in physical variability is critical for producing high-fidelity experimental data that can serve as reliable ground truth for downstream model learning and effective AI-driven experimental design.

In conclusion, the optimization of regenerative medicine manufacturing processes via CL-SDS constitutes a challenge truly worthy of the designation **Frontier**. This status is defined by its distinct **Positioning** as an unexplored white space, its profound **Significance** derived from the principle that "the process is the product," and the **Necessity** of robotic autonomy to address high dimensionality and physical sensitivity. The Robot-AI CL-SDS is, therefore, not an optional tool for efficiency; it is the **indispensable enabling methodology** required to render this complex domain tractable.

1.5 Research Challenges and Thesis Outline

Sections 1.1–1.4 established the Closed-Loop Scientific Discovery System (CL-SDS) as a unifying framework for executable, algorithmic discovery, and positioned regenerative-medicine manufacturing process development as a high-impact yet underexplored frontier. In particular, Section 1.4 argued that this frontier is defined by the principle that “the process is the product” and by a triple barrier that makes conventional human-centered process development structurally brittle: (i) High Dimensionality (combinatorially large search spaces), (ii) Time Latency (months per evaluation), and (iii) Physical Sensitivity (run-to-run and operator-to-operator variability induced by subtle physical operations). Accordingly, an effective CL-SDS for regenerative medicine can be viewed as a search-and-control system that must simultaneously navigate high-dimensional design spaces beyond human strategic planning, learn under severe wall-clock and resource budgets, and suppress experimental noise at the data-generation layer through reproducible execution. This dissertation advances CL-SDS capability in regenerative-medicine process development through two empirical studies (Chapters 2 and 3) that establish the two complementary system primitives emphasized in Chapter 4—physical loop closure and intermediate-state observability—and then integrates them into an accelerated and adaptive system perspective (Chapter 4).

1.5.1 Chapter 2: Robotic search for optimal cell culture in regenerative medicine

Chapter 2 provides the foundational proof-of-concept that CL-SDS can be physically instantiated in regenerative-medicine process development (the “White Space” identified in **Section 1.4**).

(1) Problem framing

Chapter 2 formulates iPSC-to-retinal pigment epithelium (RPE) differentiation as a high-dimensional, expensive black-box optimization problem in which the objective is to identify protocol parameter settings that improve differentiation outcome. This framing is motivated by the reality that induced differentiation is highly skill-dependent and sensitive to small physical perturbations, making manual exploration both inefficient and difficult to reproduce.

(2) System contribution

Chapter 2 implements a physically closed CL-SDS by integrating: (i) batch Bayesian optimization for experiment design, (ii) a dual-arm humanoid laboratory robot (Maholo) for

repeatable execution at scale, and (iii) an automated scoring pipeline that converts imaging data into quantitative feedback to construct a Bayesian surrogate model. The system is designed around the practical constraint of parallel execution (48 wells per round) and incorporates hardware-dependent context (e.g., structured constraints on trypsin timing) into the batch query generation policy.

(3) Evaluation strategy

As the true end goal of cell manufacturing quality is multi-factorial and cannot be exhaustively evaluated at every iteration, Chapter 2 adopts a practical surrogate objective—the pigmentation score—as the operational readout for closed-loop search. This choice is treated explicitly as an engineering proxy: it enables standardized, relatively high-throughput feedback for systematic exploration while downstream biological assays are used to confirm that optimized conditions yield RPE cells that satisfy key characteristics expected for regenerative-medicine research use.

(4) Key result and why it matters

Across a search space of ~200 million parameter combinations, the system executed 143 conditions over 111 days (three iterative rounds of 48-well parallel experiments in a seven-variable parameterization) and achieved a substantial improvement relative to pre-optimized culture in terms of pigmentation score (reported as 88% improvement in the chapter summary), followed by biological validation of the resulting RPE cells. The core scientific significance is not merely “automation,” but the demonstration that regenerative-medicine process development—traditionally dominated by tacit skills and uncontrolled variability—can be reframed as an executable CL-SDS loop in which robotics suppresses physical variance while AI performs strategic exploration in high dimensions.

1.5.2 Chapter 3: Early and non-destructive prediction of differentiation efficiency using imaging and machine learning

While Chapter 2 successfully demonstrated the physical closure of the CL-SDS loop using robotics and AI, a critical challenge remains: the temporal latency of feedback. In many regenerative protocols, the terminal phenotype defining success appears only after extended culture periods. This latency renders the optimization cycle slow if the system must wait for the final readout at every iteration. Chapter 3 addresses this “time barrier” by establishing a methodology for early, non-destructive outcome prediction and thereby making long culture trajectories partially observable.

(1) Problem framing

Chapter 3 targets the time/latency barrier that remains even after physical loop closure: in many regenerative protocols, the terminal readout that defines success appears only after long induction periods. The case study is a directed differentiation protocol from hiPSCs to muscle stem cells (MuSCs), where terminal induction efficiency is assessed by MYF5 positivity on approximately day 82. In such settings, naive closed-loop optimization becomes unacceptably slow because each iteration consumes months before objective feedback is available.

(2) System contribution

Chapter 3 develops an early, non-destructive prediction module using phase-contrast imaging and machine learning to infer terminal induction efficiency from intermediate-stage morphology. Given the intrinsically limited number of feasible long-protocol experiments (34 wells across six independent experiments), the work deliberately adopts a robust classical machine learning pipeline rather than data-hungry deep learning, with careful evaluation to mitigate overfitting and leakage risk.

(3) Evaluation strategy

The study first identifies biologically plausible windows for prediction (the myogenic induction phase), then validates predictability through day-by-day classification tasks that discriminate high versus low terminal MYF5+% groups, using grouped cross-validation and repeated trials to quantify stability. In addition to standard classifier metrics (AUC), the model is evaluated in an operationally meaningful “filter” scenario, quantifying how early prediction could enrich high-quality samples or remove defective ones without destructive assays.

(4) Key result and why it matters

Chapter 3 shows that terminal MYF5+% outcomes at day 82 are predictable more than 50 days in advance, with particularly informative time points around day 24 (for low-efficiency detection) and days 31–34 (for high-efficiency detection), achieving AUCs on the order of ~0.76–0.86 depending on task and day. Notably, whereas prior imaging-based forecasting studies in stem-cell differentiation typically focused on induction periods shorter than ~1 month and produced predictions only within days of the terminal readout, this work demonstrates actionable predictability at an early phase of an ~82-day directed differentiation protocol. Conceptually, this result establishes that long-horizon differentiation trajectories contain early, non-destructive signatures correlated with terminal fate—an enabling condition for shortening effective feedback cycles via surrogate evaluation, early stopping, or multi-fidelity closed-loop strategies in regenerative-medicine CL-SDS.

1.5.3 Chapter 4: General discussion and Conclusion

Chapters 2 and 3 jointly deliver two system primitives required to operationalize CL-SDS in regenerative-medicine manufacturing: (i) cyber-physical loop closure with reduced execution variability (Chapter 2), and (ii) intermediate-state observability via early, non-destructive outcome prediction (Chapter 3). Building on these results, Chapter 4 articulates how CL-SDS can evolve from “optimizing a static protocol with a long terminal readout” into accelerated and adaptive architectures that exploit intermediate information to improve both cycle time and robustness. Concretely, Chapter 4 develops and motivates two next-step directions:

(1) Accelerated discovery through multi-fidelity closed-loop optimization

With an early predictor available, the discovery loop can be redesigned as a multi-fidelity closed-loop optimizer: intermediate predictions are treated as low-cost, low-fidelity “weak observations” that can be injected into the planner to enable earlier triage, earlier model updates, and earlier proposal of new conditions, while terminal assays remain the high-fidelity ground truth used for confirmation and calibration. In practice, this perspective naturally supports asynchronous/parallel planning—allowing subsequent

experimental rounds to be designed before all terminal readouts are available—thereby reducing wall-clock time per learning cycle without conflating proxies with final quality.

(2) Reducing quality variability via state estimation & feedback control

Intermediate-state estimation also enables a shift from static, open-loop protocols to adaptive control: if a culture trajectory deviates from the region associated with successful outcomes, corrective actions may be applied mid-course. This direction is particularly important for regenerative medicine, especially autologous and patient-specific products, where “initial-condition uncertainty” in starting material and unavoidable lot-to-lot or patient-to-patient variability can propagate through long culture trajectories and appear as batch-to-batch fluctuation in final CQAs (or outright batch failure) even when execution is standardized. An observation-driven control layer offers a principled pathway to reducing quality dispersion and improving yield and robustness under biological uncertainty.

1.5.4 Thesis structure

In summary, Chapter 2 establishes the core feasibility of CL-SDS in regenerative-medicine process development by implementing an autonomous robotic closed loop for protocol parameter search and demonstrating meaningful performance gains under a practical surrogate objective. Chapter 3 adds a complementary module—early, non-destructive prediction from imaging—that directly targets the time-latency barrier by making long culture trajectories partially observable and enabling time-axis compression of feedback. Finally, Chapter 4 integrates these two primitives into an architectural discussion of how CL-SDS in regenerative medicine can evolve toward multi-fidelity acceleration and trajectory-level feedback control, outlining the next technical milestones needed to operationalize accelerated and adaptive CL-SDS.

References

1. Bacon, S. F. & Devey, J. Novum Organum. https://oll.libertyfund.org/titles/bacon-novum-organum.
2. Popper, K. R. The logic of scientific discovery. *pp* **480**, (1959).
3. Baker, M. 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
4. Kanie, K. *et al.* Effect of mechanical vibration stress in cell culture on human induced pluripotent stem cells. *Regen Ther* **12**, 27–35 (2019).
5. Holland, I. & Davies, J. A. Automation in the life science research laboratory. *Front. Bioeng. Biotechnol.* **8**, 571777 (2020).
6. Kitano, H. Artificial Intelligence to Win the Nobel Prize and Beyond: Creating the Engine for Scientific Discovery. *AIMag* **37**, 39 (2016).

7. Wilson, J. T., Hutter, F. & Deisenroth, M. P. Maximizing acquisition functions for Bayesian optimization. *arXiv [stat.ML]* (2018).
8. King, R. D. *et al.* The automation of science. *Science* **324**, 85–89 (2009).
9. Fisher, R. A. The design of experiments. *xi* **251**, (1935).
10. Box, G. E. P. & Wilson, K. B. On the experimental attainment of optimum conditions. *J. R. Stat. Soc. Series B Stat. Methodol.* **13**, 1–38 (1951).
11. Buchanan, B. G. & Feigenbaum, E. A. Dendral and meta-dendral: Their applications dimension. *Artif. Intell.* **11**, 5–24 (1978).
12. W., P. Bacon: A production system that discovers empirical laws. https://www.ijcai.org/Proceedings/77-1/Papers/057.pdf.
13. Schmidt, M. & Lipson, H. Distilling free-form natural laws from experimental data. *Science* **324**, 81–85 (2009).
14. Newell, A. & Simon, H. A. Human problem solving. *xiv* **920**, (1972).
15. Vicini, A. & Quagliarella, D. Inverse and direct airfoil design using a multiobjective genetic algorithm. *AIAA J.* **35**, 1499–1505 (1997).
16. Koza, J. R., Bennett, F. H., III, Andre, D. & Keane, M. A. Automated design of both the topology and sizing of analog electrical circuits using genetic programming. in *Artificial Intelligence in Design '96* 151–170 (Springer Netherlands, Dordrecht, 1996). doi:10.1007/978-94-009-0279-4_9.
17. Willett, P. Genetic algorithms in molecular recognition and design. *Trends Biotechnol.* **13**, 516–521 (1995).
18. Močkus, J. On Bayesian methods for seeking the extremum. in *Optimization Techniques IFIP Technical Conference* 400–404 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1975). doi:10.1007/978-3-662-38527-2_55.
19. Jones, D. R., Schonlau, M. & Welch, W. J. Efficient Global Optimization of Expensive Black-Box Functions. *J. Glob. Optim.* **13**, 455–492 (1998).
20. Skeggs, L. T., Jr. An automatic method for colorimetric analysis. *Am. J. Clin. Pathol.* **28**, 311–322 (1957).

21. Merrifield, R. B. Solid phase peptide synthesis. I. the synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
22. Little, J. N. The Zymate Laboratory Automation Systems. *J. Liq. Chromatogr.* **9**, 3197–3201 (1986).
23. Pereira, D. A. & Williams, J. A. Origin and evolution of high throughput screening. *Br. J. Pharmacol.* **152**, 53–61 (2007).
24. Lin, B. B. High throughput screening for new drug discovery. *J. Food Drug Anal.* **3**, (1995).
25. Nikolaev, P. *et al.* Autonomy in materials research: a case study in carbon nanotube growth. *Npj Comput. Mater.* **2**, 16031 (2016).
26. Burger, B. *et al.* A mobile robotic chemist. *Nature* **583**, 237–241 (2020).
27. MacLeod, B. P. *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Sci. Adv.* **6**, eaaz8867 (2020).
28. Steiner, S. *et al.* Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **363**, eaav2211 (2019).
29. Mehr, S. H. M., Craven, M., Leonov, A. I., Keenan, G. & Cronin, L. A universal system for digitization and automatic execution of the chemical synthesis literature. *Science* **370**, 101–108 (2020).
30. Coley, C. W. *et al.* A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, eaax1566 (2019).
31. Dave, A. *et al.* Autonomous optimization of non-aqueous Li-ion battery electrolytes via robotic experimentation and machine learning coupling. *Nat. Commun.* **13**, 5454 (2022).
32. Williams, K. *et al.* Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases. *J. R. Soc. Interface* **12**, 20141289 (2015).
33. Parry, D. M. Closing the loop: Developing an integrated design, make, and test platform for discovery. *ACS Med. Chem. Lett.* **10**, 848–856 (2019).
34. Brocklehurst, C. E. *et al.* MicroCycle: An integrated and automated platform to accelerate drug discovery. *J. Med. Chem.* **67**, 2118–2128 (2024).

35. Grace, J. *et al.* Engineering multispecific antibodies with complete killing selectivity through the closed-loop integration of machine learning and high-throughput experimentation. *MAbs* **17**, 2598093 (2025).
36. Zhavoronkov, A. *et al.* Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* **37**, 1038–1040 (2019).
37. DeMeo, B. *et al.* Active learning framework leveraging transcriptomics identifies modulators of disease phenotypes. *Science* **390**, eadi8577 (2025).
38. Ha, T. *et al.* AI-driven robotic chemist for autonomous synthesis of organic molecules. *Sci. Adv.* **9**, eadj0461 (2023).
39. Chao, R., Mishra, S., Si, T. & Zhao, H. Engineering biological systems using automated biofoundries. *Metab. Eng.* **42**, 98–108 (2017).
40. Holowko, M. B., Frow, E. K., Reid, J. C., Rourke, M. & Vickers, C. E. Building a biofoundry. *Synth. Biol. (Oxf.)* **6**, ysaa026 (2021).
41. HamediRad, M. *et al.* Towards a fully automated algorithm driven platform for biosystems design. *Nat. Commun.* **10**, 5150 (2019).
42. Hillson, N. *et al.* Author Correction: Building a global alliance of biofoundries. *Nat. Commun.* **10**, 3132 (2019).
43. Bates, M., Berliner, A. J., Lachoff, J., Jaschke, P. R. & Groban, E. S. Wet Lab Accelerator: A web-based application democratizing laboratory automation for synthetic biology. *ACS Synth. Biol.* **6**, 167–171 (2017).
44. Miles, B. & Lee, P. L. Achieving reproducibility and closed-loop automation in biological experimentation with an IoT-enabled lab of the future. *SLAS Technol.* **23**, 432–439 (2018).
45. Adam, D. The automated lab of tomorrow. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2406320121 (2024).
46. Hitachi, L. Hitachi's new cell design platform, which integrates digital and biotechnology, will streamline drug discovery for cell and gene therapy. https://www.hitachi.com/New/cnews/month/2025/05/250513b.html.

47. Kino-Oka, M. *et al.* Automating the expansion process of human skeletal muscle myoblasts with suppression of myotube formation. *Tissue Eng. Part C Methods* **15**, 717–728 (2009).
48. Nishimura, A. *et al.* Fabrication of tissue-engineered cell sheets by automated cell culture equipment. *J. Tissue Eng. Regen. Med.* **13**, 2246–2255 (2019).
49. dos Santos, F. F., Andrade, P. Z., da Silva, C. L. & Cabral, J. M. S. Bioreactor design for clinical-grade expansion of stem cells. *Biotechnol. J.* **8**, 644–654 (2013).
50. Konagaya, S., Ando, T., Yamauchi, T., Suemori, H. & Iwata, H. Long-term maintenance of human induced pluripotent stem cells by automated cell culture system. *Sci. Rep.* **5**, 16647 (2015).
51. Thomas, R. J., Chandra, A., Hourd, P. C. & Williams, D. J. Cell Culture Automation and Quality Engineering: A Necessary Partnership to Develop Optimized Manufacturing Processes for Cell-Based Therapies. *JALA: Journal of the Association for Laboratory Automation* **13**, 152–158 (2008).
52. Liu, Y., Hourd, P., Chandra, A. & Williams, D. J. Human cell culture process capability: a comparison of manual and automated production. *J. Tissue Eng. Regen. Med.* **4**, 45–54 (2010).
53. Soares, F. A. C. *et al.* Investigating the feasibility of scale up and automation of human induced pluripotent stem cells cultured in aggregates in feeder free conditions. *J. Biotechnol.* **173**, 53–58 (2014).
54. Matsumoto, E. *et al.* Fabricating retinal pigment epithelial cell sheets derived from human induced pluripotent stem cells in an automated closed culture system for regenerative medicine. *PLoS One* **14**, e0212369 (2019).
55. Ochiai, K. *et al.* A Variable Scheduling Maintenance Culture Platform for Mammalian Cells. *SLAS Technol* **26**, 209–217 (2021).
56. Mason, C. & Hoare, M. Regenerative medicine bioprocessing: building a conceptual framework based on early studies. *Tissue Eng.* **13**, 301–311 (2007).
57. Lipsitz, Y. Y., Timmins, N. E. & Zandstra, P. W. Quality cell therapy manufacturing by

design. *Nat. Biotechnol.* **34**, 393–400 (2016).

58. *Manufacturing Changes and Comparability for Human Cellular and Gene Therapy Products: Draft Guidance for Industry*.
59. Ayala Ceja, M., Khericha, M., Harris, C. M., Puig-Saus, C. & Chen, Y. Y. CAR-T cell manufacturing: Major process parameters and next-generation strategies. *J. Exp. Med.* **221**, (2024).
60. Holm, P., Allesø, M., Bryder, M. C. & Holm, R. Q8(R2). in *ICH Quality Guidelines* 535–577 (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2017). doi:10.1002/9781118971147.ch20.
61. Raghavan, R. R. & McCombie, R. ICH Q5E comparability of biotechnological/biological products subject to changes in their manufacturing processes. in *ICH Quality Guidelines* 409–432 (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2017). doi:10.1002/9781118971147.ch15.
62. Feldman, R. Trade secrets and biologic medicine: A conflict of laws. *SSRN Electron. J.* (2022) doi:10.2139/ssrn.4230581.

Chapter2

2. Robotic search for optimal cell culture in regenerative medicine

2.1 Chapter Summary

Induced differentiation is one of the most experience- and skill-dependent experimental processes in regenerative medicine, and establishing optimal conditions often takes years. I developed a robotic AI system with a batch Bayesian optimization algorithm that autonomously induces the differentiation of induced pluripotent stem cell-derived retinal pigment epithelial (iPSC-RPE) cells. From 200 million possible parameter combinations, the system performed cell culture in 143 different conditions in 111 days, resulting in 88% better iPSC-RPE production than that obtained by the pre-optimized culture in terms of the pigmentation scores. Our work demonstrates that the use of autonomous robotic AI systems drastically accelerates systematic and unbiased exploration of experimental search space, suggesting immense use in medicine and research.

2.2 Note on Contribution

Paper information: The content of this chapter has been published in the following journal article:

Genki N Kanda*, **Taku Tsuzuki***, Motoki Terada, Noriko Sakai, Naohiro Motozawa, Tomohiro Masuda, Mitsuhiro Nishida, Chihaya T Watanabe, Tatsuki Higashi, Shuhei A Horiguchi, Taku Kudo, Motohisa Kamei, Genshiro A Sunagawa, Kenji Matsukuma, Takeshi Sakurada, Yosuke Ozawa, Masayo Takahashi, Koichi Takahashi, Tohru Natsume (2022) Robotic search for optimal cell culture in regenerative medicine eLife 11:e77007 (* The authors contributed equally to this work)

Clarification of Contributions: This study was conducted as an interdisciplinary collaborative project integrating biology, robotics, and artificial intelligence. As a co-first author, I was primarily responsible for the computational framework, mathematical problem formulation, and software system architecture. Specifically, my contributions are as follows:

1. Conceptualization and Problem Definition

Core Concept Proposition: I proposed the core idea of the study: that the optimization of iPSC differentiation protocols—typically plagued by variability in manual techniques—could be formulated and solved as a mathematical optimization problem by leveraging the high reproducibility of robotic systems. This concept was conceived with the aim of achieving a world-first demonstration of AI-driven optimization in regenerative medicine.
Problem Formulation: I redefined the biological process of iPSC-RPE differentiation as a "black-box optimization problem." Based on domain knowledge and preliminary simulations, I defined their search spaces to ensure the optimization was computationally feasible.

2. Algorithm Development and Implementation

Bayesian Optimization Engine: I designed and developed the experimental design engine using Bayesian optimization. To address the constraint of conducting 48 parallel experiments with hardware-specific batch limitations, I developed and implemented the Batch Contextual Local Penalization (BCLP) algorithm.
Image Processing Pipeline: I designed and implemented the automated image processing module that converts microscopic phase-contrast images into quantitative "pigmentation scores."

3. Software System Integration

I designed the software and built a compiler that automatically translates the experimental parameters generated by the AI engine into executable robot instruction sets, realizing the "closed-loop" automation system.

4. Data Analysis and Manuscript Writing

I performed the statistical analysis of the optimization results and visualized the convergence of parameters. I drafted the manuscript, including the Introduction and Discussion, in addition to the technical sections on computational methodology and system design.

Contributions of Co-authors: The co-first author, Genki N. Kanda, and the biological/robotics teams (RIKEN and Robotic Biology Institute) were primarily responsible for the robotization of the baseline protocol, wet-lab experiments (e.g., cell maintenance, reagent preparation, robot operations), and biological validation (e.g., PCR, ELISA). They also contributed to the overall conceptualization of the project and the writing of the manuscript.

2.3 Introduction

Automating scientific discovery is one of the grandest challenges of the 21st century[1,2]. A promising approach involves creating a closed loop of computation and experimentation by combining AI and robotics [3]. A relatively simple form of autonomous knowledge discovery involves searching for optimal experimental procedures and parameter sets through repeated experimentation and result validation, according to a predefined validation method. For example, in material science, the parameters associated with the growth of carbon nanotubes have been explored using an autonomous closed-loop learning system [4]. In experimental physics, Bayesian optimization has been used to identify the optimal evaporation ramp conditions for Bose–Einstein condensate production [5]. In 2019, a promoter-combination search in molecular biology was automated using an optimization algorithm-driven robotic system [6]. Some robotic systems for cell culture have already been developed [7–15]; however, many of these fixed-process automation apparatuses lack the flexibility and precision necessary to execute comprehensive parameter searching.

Here, I report the development of a robotic search system that autonomously and efficiently searches for the optimal conditions for inducing iPS cell differentiation into retinal pigment epithelial (RPE) cells (iPSC-RPE cells). The system replaces the manual operations involved in cell culture with robotic arms. Cell culture is probably one of the most delicate

procedures in two respects. First, the parameters related to physical manipulation can greatly affect the outcome of the experiment [16]. Secondly, it takes a long time to execute a series of protocols. For example, cells artificially differentiated from embryonic stem cells or induced pluripotent stem cells (ES/iPS cells) need to be processed using hundreds of experimental procedures that typically last for weeks or months before they can be used for transplantation in regenerative medicine.

During these processes, cells are given chemical perturbations (e.g., type, dose, and timing of reagents) and physical perturbations (e.g., strength of pipetting, vibration during handling of plates, timing of transfer from/to CO_2 incubator, and accompanying changes in factors such as temperature, humidity, and CO_2 concentration). Due to the heterogeneous and complex internal states of cells, suitable culture conditions must be determined for each strain and/or lot [17]. A small difference in a single chemical stimulus or physical procedure can lead to failure of differentiation or poor quality of the produced cells, and such consequences can often become experimentally detectable only days or weeks after the input is given [18]. Therefore, the use of robotic arms is a great addition in the search for optimal cell culture conditions because robots can repeatedly perform the same operation with high precision. Moreover, they hardly make any errors, which are logged when committed.

It is advantageous to utilize high-accuracy and programmable robotic arms for the search of optimal cell culture parameters. Unlike human hands, robotic arms can repeatedly perform the same procedure. They ensure reproducibility by keeping all parameters related to physical procedures constant. Furthermore, the actual operations are logged by the software along with sensor information when they are deviated from the established programs. Thus, robotization provides an ideal parameterization of experimental procedures. Some automated cell culture machines have already been proposed [19]; however, proper formulation of an autonomous search for optimal culture conditions has not yet been determined.

In this study, I combined a Maholo LabDroid [20] and an AI system that independently evaluates the experimental results and plans the next experiments to realize an autonomous robotic search for optimal culture conditions. I first created a digital representation of the regenerative medical cell culture protocol used for iPS cell differentiation into retinal pigment epithelial (RPE) cells (iPSC-RPE cells) [21], which can be executed by the robot and used as a template for an AI-driven parameter search. I then implemented the experimental protocol on a LabDroid, which is a versatile humanoid robot that can perform a broad range of experimental procedures. Its flexibility allows frequent changes in protocols and protocol parameters, making it suitable for use in experimental parameter searches. The robot has an integrated microscope that provides data for image-processing through AI, which evaluates the quality of growing cells. The search process was mathematically formulated as a type of experimental design problem, and a batch Bayesian optimization (BBO; **Figure 2.1, Figure 2.2**) technique was employed as a solver. Finally, I demonstrated that iPSC-RPE cells generated by LabDroid satisfy the cell biological criteria for regenerative medicine research applications.

The architecture of our autonomous robotic search system is illustrated in **Figure 2.2**. The system functions as a closed-loop platform integrating computational planning with physical execution. The workflow begins with the Optimizer generating a set of experimental parameters (queries), which are transmitted to the Protocol Compiler. The compiler translates these abstract parameters into specific machine instructions executable by the LabDroid, combined with the base experimental protocol defined by the user (**Figure 2.17-21**). Inside the LabDroid booth, the dual-arm robot performs the physical wet-lab

operations—handling reagents, pipetting, and managing cell cultures in the incubator—as indicated by the dotted lines in the figure. Following the culture period, the integrated Microscope captures phase-contrast images of the cells. These images are processed by the Scoring process module to quantify the differentiation efficiency (pigmentation score) (**Figure 2.22**). Finally, this score is fed back to the Optimizer to update the probabilistic model and generate the next set of queries, completing the autonomous cycle without human intervention(**Figure 2.10, Algorithm1-3**).

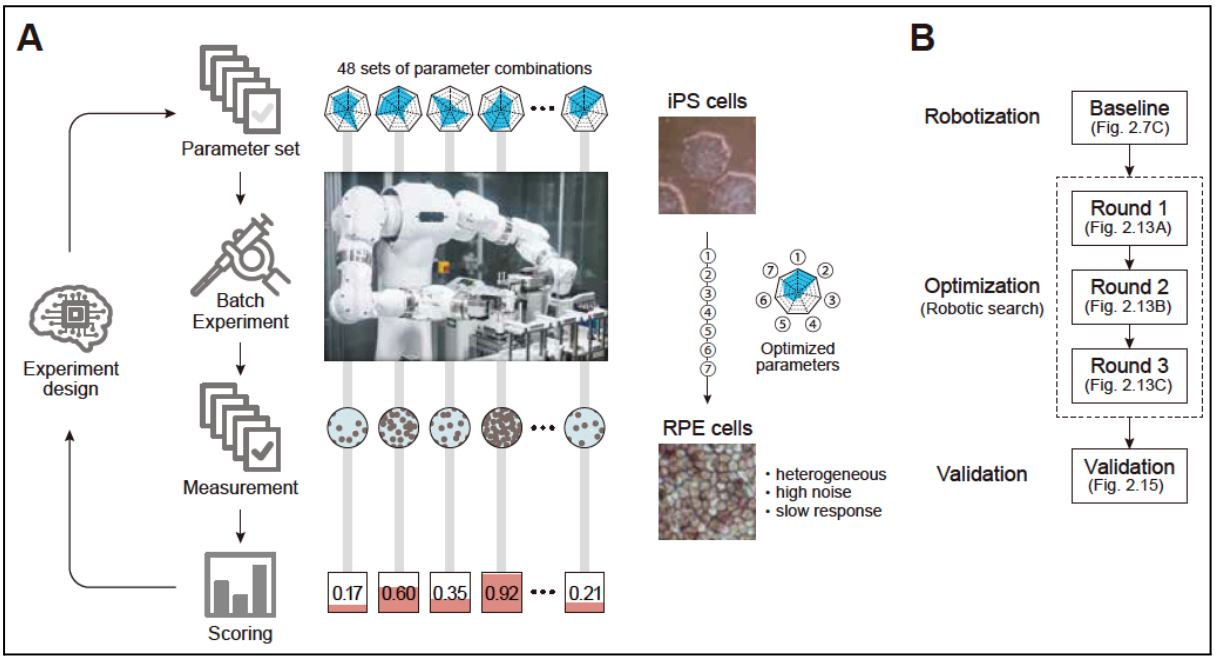


Figure 2.1: Robotic search for optimal experimental conditions.
(**A**) Overall workflow for the optimization of experimental procedures using combined experimental robotics and Bayesian optimization. The user defines the target experimental protocol, subject parameters of the protocol, and the validation function. In this study, I chose the differentiation procedure from iPS to RPE cells as a target protocol and selected the reagent concentration, administration period, and five other parameters (details are shown in **Table 1**). I defined the pigmented area in a culture well, which represents the degree of RPE differentiation induction, as the validation function. The optimization program presented multiple parameter candidates; the LabDroid performed the experiment, and then an evaluation value for each candidate was obtained. Subsequently, the Bayesian optimization presented a plurality of parameter candidates predicted to produce higher validation values. The optimal parameters were searched by repeating candidate presentation, experiment execution, validation, and prediction. The detailed components are shown in **Figure 2.2**.
(**B**) Workflows performed in this study. First, robotization of the iPSC-RPE protocol was performed as a baseline. Next, the optimization process was conducted in three rounds, followed by statistical and biological validation. The figure numbers in parentheses represent the results shown in the figure.

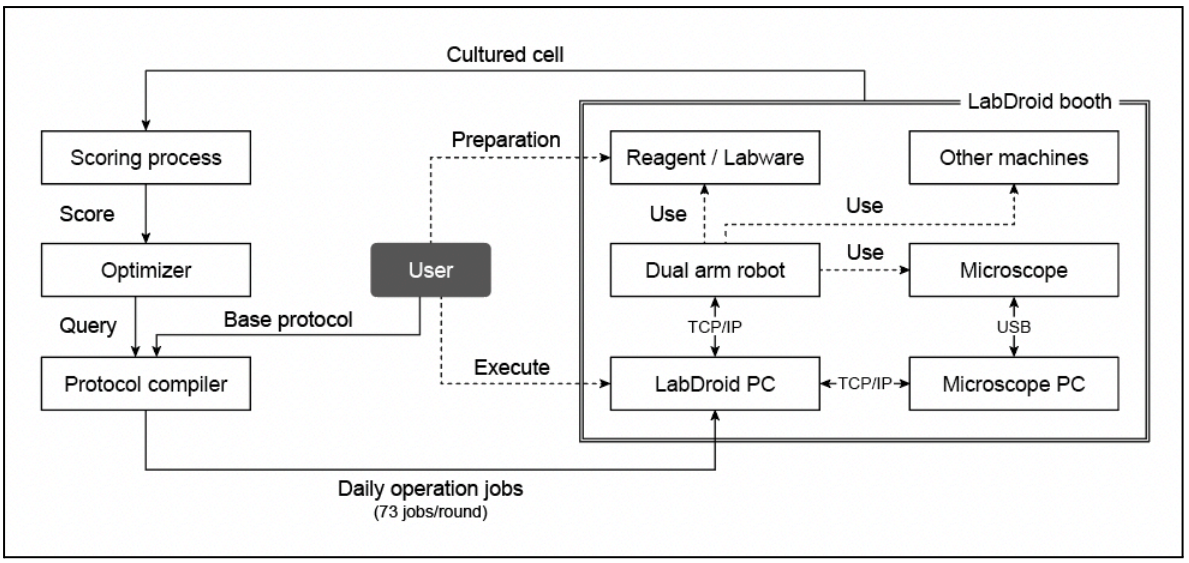


Figure 2.2: System components

The system platform. The robotic search using this system starts with the provision of queries for round 1 from the optimizer to the protocol compiler and the provision of the base protocol from the user to the protocol compiler. Each rectangle represents a system component. Solid lines represent the movement of data (including cells), and dotted lines represent physical interactions.

2.4. Results

2.4.1. Robotization of the iPSC-RPE differentiation protocol

Overview of the iPSC-RPE differentiation protocol

To contextualize the scope of our robotic optimization, **Figure 2.3** outlines the comprehensive workflow required for generating iPSC-derived RPE cells for clinical transplantation . The entire process is extensive, spanning from the initial iPSC expansion and differentiation induction to purification, storage, and final recovery culture prior to surgery. While the clinical application requires strict quality control across all these stages, our study specifically targets the **iPSC-RPE differentiation phase** (highlighted in yellow in **Figure 2.3**). This phase typically involves complex manual operations over a period of 5–6 weeks and is a major source of variability. By isolating this critical window, I aimed to demonstrate that robotic optimization can replace and improve upon one of the most experience-dependent segments of the regenerative medicine pipeline.

An overview of the iPSC-RPE differentiation protocol used for optimization is shown in **Figure 2.4**. It consists of five steps: seeding, preconditioning, passage, RPE differentiation (induction), and RPE maintenance culture. The day on which the passage was performed was defined as differentiation day (DDay) 0, and the cultured cells were sampled and validated on DDays 33 and 34.

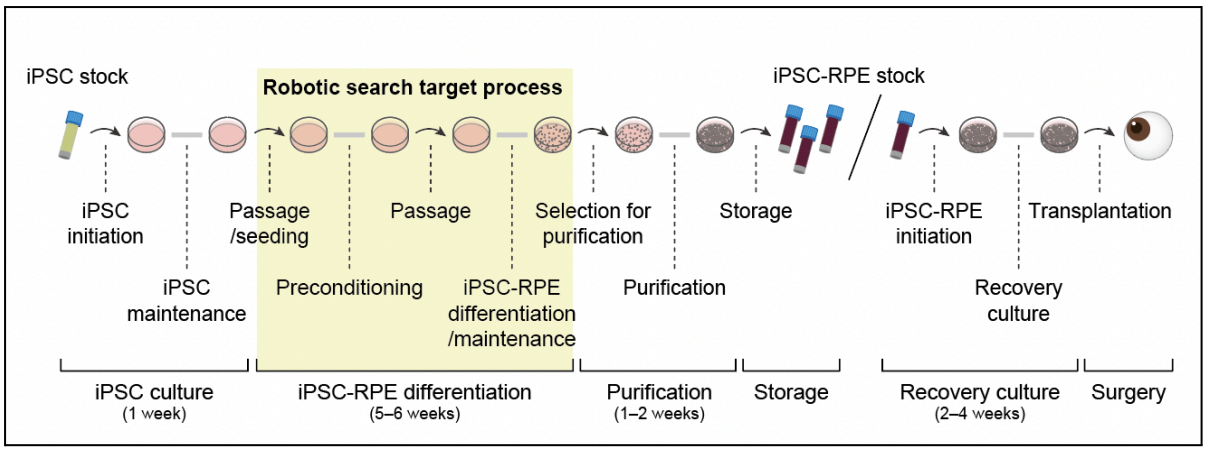


Figure 2.3: Process flow from iPSC stock to iPSC-derived RPE cell transplantation
Schematic diagram of the process from iPSC stock to iPSC-derived RPE cell transplantation. The steps are roughly divided into iPSC culture, iPSC-RPE differentiation, purification, storage, recovery culture, and surgery. Arrows indicate daily operations, and thick gray lines indicate multi-day operations.

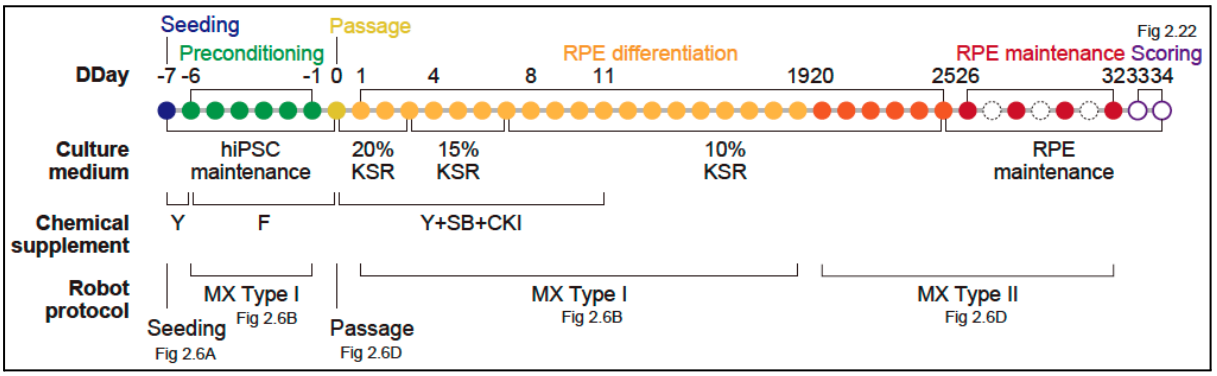


Figure 2.4: Timeline of the robotic iPSC-RPE differentiation procedure
Schematic diagram of the standard iPSC-RPE differentiation procedures. DDay indicates the differentiation day. Filled circles represent days when the robot operated, solid circles represent days with human operations only, and dashed line circles represent days when no operations were conducted. F stands for FGF receptor inhibitor; Y for Y-27632, a Rho-kinase inhibitor; SB for SB431542, a TGF-β/Activin/Nodal signal inhibitor; CK for a CKI-7, Wnt signal inhibitor; and MX for medium exchange.

Overview of the LabDroid's workbench

To implement this protocol using LabDroid, the necessary peripheral devices were installed on and around LabDroid's workbench (**Figure 2.5**, **Figure 2.16**). I designed the system to work simultaneously with eight 6-well plates per batch, for a total of 48 cell-containing wells. LabDroid was programmed for three types of operations: seeding, medium exchange, and passage (**Figure 2.5, Figure 2.17-21**). The steps for the preconditioning and induction, which correspond to the preparation of reagents, were named medium exchange type I, and the step for RPE maintenance culture, which does not involve reagent preparation, was named medium exchange type II (**Figure 2.4**, **Figure 2.6**).

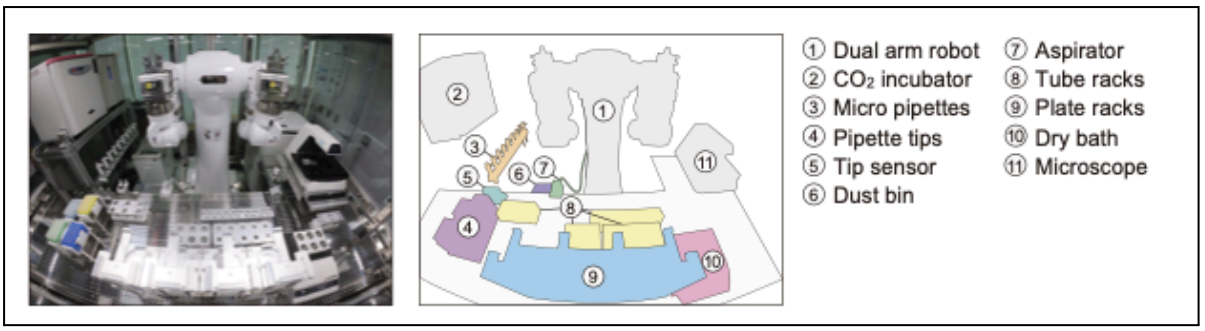


Figure 2.5: Layout of peripheral devices on the LabDroid's workbench.
The LabDroid Maholo including peripheral equipment.
Components: (1) dual-arm humanoid; (2) CO_2 incubator; (3) micropipettes; (4) pipette tips; (5) tip sensor; (6) dustbin; (7) aspirator; (8) tube racks; (9) plate racks; (10) dry bath; and (11) microscope.

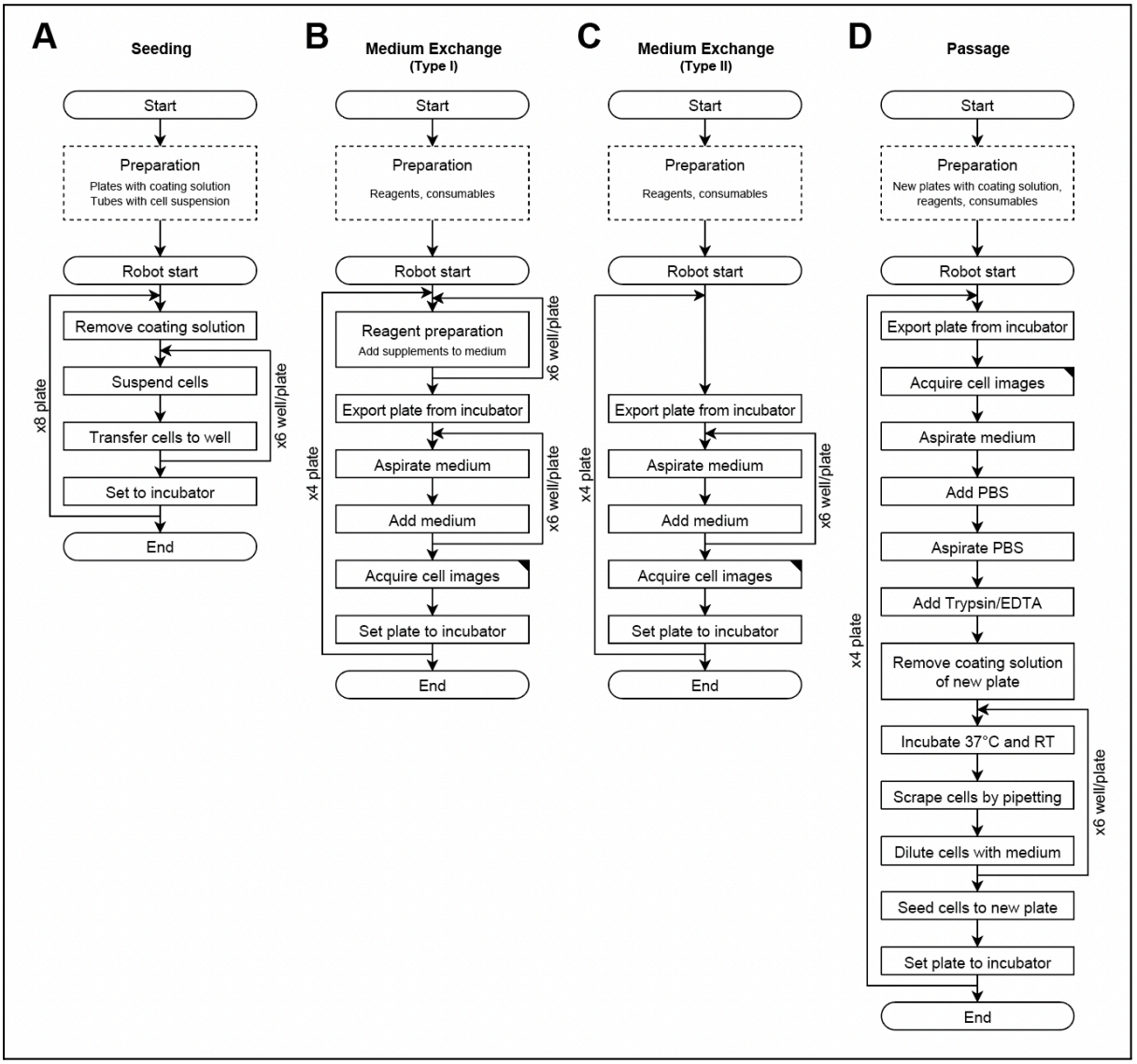


Figure 2.6: Workflows of the experimental operation.
(**A**) Seeding (DDay −7).

(**B**) Medium exchange type I for preconditioning (DDays −6 to −1) and the first part of RPE induction (DDays 1 to 19).
(**C**) Medium exchange type II for the second part of RPE induction (DDays 20 to 25), and RPE maintenance (DDays 26 to 32).
(**D**) Passage (DDay 0).
The dashed line rectangles represent the operations carried out by humans, and solid line rectangles represent the operations carried out by the robot.

Baseline experiments

First, I used LabDroid to perform baseline experiments involving the induction of iPSC-RPE cell differentiation under the same conditions as the typical manual operations. Because of the differences in structure and experimental environment between the LabDroid and humans, some operations and movements, such as the use of a centrifuge, the presence or absence of cell counting at the time of passage, and the speed of movement, differed from those of humans. For example, achieving the same time interval for trypsin treatment in all wells of a single plate during cell detachment using LabDroid is difficult. Therefore, the passage operation was performed at six separate time intervals. The cells differentiating into RPE cells produce melanin, which causes them to turn brown. Therefore, the area ratio of the total number of pigmented cells on DDay 34 was used to estimate the differentiation induction efficiency and obtain evaluation scores, following the example of previous studies [19,22] (**Figure 2.22**). These validation scores were used to simplify the validation process and do not reflect the entire quality of the RPE.

Baseline experiments were conducted and validated using six trypsin conditions and eight plates (**Figure 2.7**, **Figure 2.8**). The highest scoring was obtained when trypsin treatment was conducted for 20 min at 37 °C, followed by 14 min incubation at room temperature (RT, approximately 25 °C), with an eight-plate score of 0.44 ± 0.03 (mean ± SEM, n = 8). The lowest scoring was obtained when trypsin treatment was conducted for 20 min at 37 °C, followed by 23 min at RT, with an eight-plate score of 0.33 ± 0.02 (mean ± SEM, n = 8). LabDroid successfully performed the iPSC-RPE protocol, as evidenced by the detection of pigmented cells in all 48 wells and the lack of errors in the operating process. However, in the naive transplantation of the manual protocol to the robot, the induction efficiency was insufficient. This suggests that it is inherently difficult to describe physical parameters, including unrecorded human movements. Therefore, I attempted to optimize the protocol parameters to further improve the scores using a robotic search.

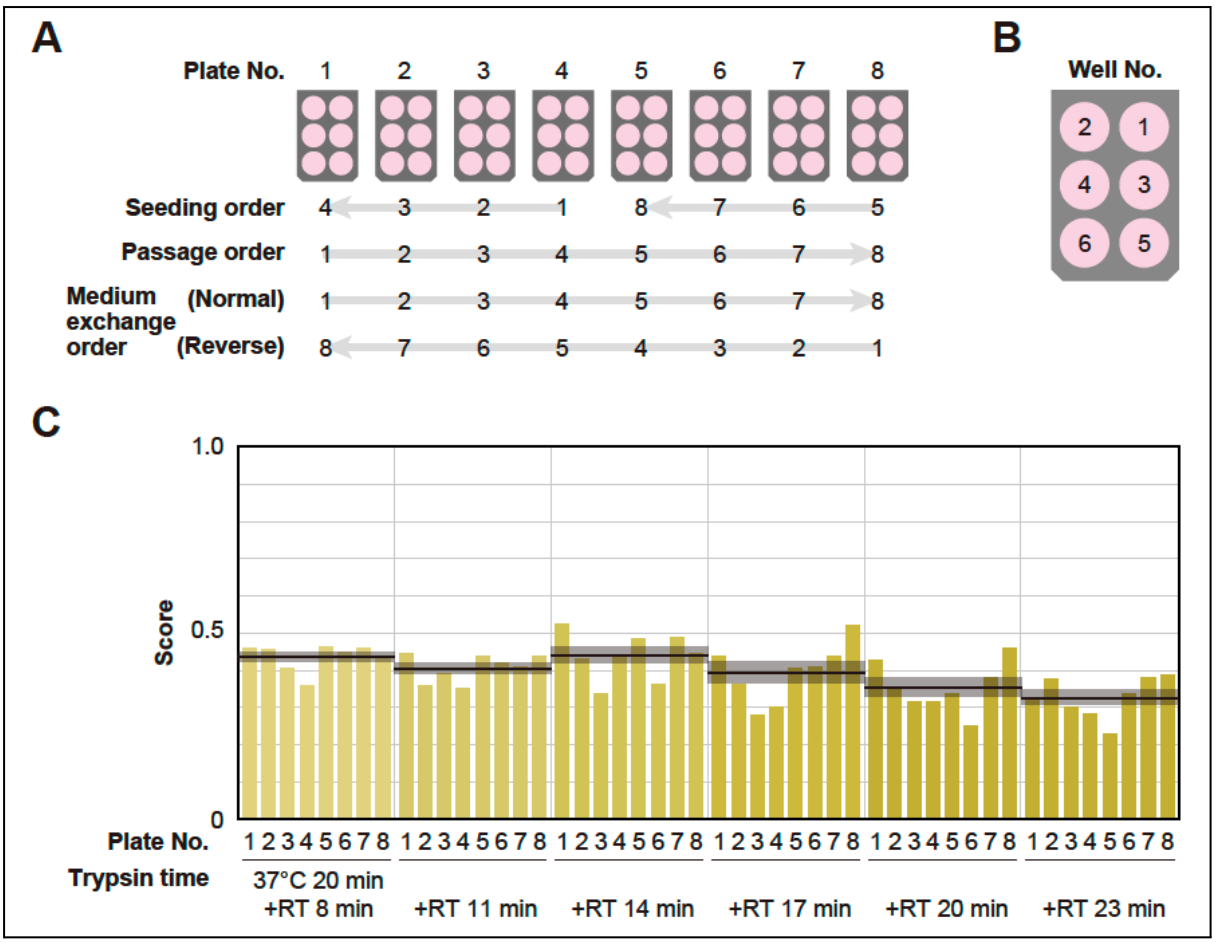


Figure 2.7: Baseline evaluation of the robotic iPSC-RPE differentiation.
(**A**) Plate numbering and the orders of seeding, passage, and medium exchange operations. Eight 6-well plates were used for each experiment.
(**B**) Well numbering.
(**C**) Scores of the first trial. iPSC-RPE differentiation was conducted under six different trypsin treatment times using the LabDroid. Yellow bars represent the pigmented cell area score of each well. The bold black lines and the shaded area around the lines represent the mean score and SEM of eight samples operated at the same trypsin time, respectively.



Figure 2.8: Acquired pigmented images of the baseline experiment.
Images acquired on Day 34 of the baseline experiment; images of the bottom of the well with cultured cells, cropped to the size of the well. These 8-bit images were adjusted to a minimum and maximum contrast value of 100 and 150, respectively. IDs on the bottom indicate 'B (baseline) - Plate No. - Well No.'.

2.4.2. Parameterization of the protocol

To improve the pigmentation score, I selected seven parameters for optimization: two from the preconditioning step, three from the passage step, and two from the induction step. Search domains were set for each parameter (**Table 2.1**; **Figure 2.9**).

From the preconditioning step on DDays −1 to −6, I selected two parameters for optimization: the concentration of fibroblast growth factor receptor inhibitor (FGFRi) in the medium (*PC*, preconditioning concentration), and the duration of addition (*PP*, preconditioning period). From the passage step performed on DDay 0, I selected three parameters to optimize: the pipetting strength during cell detachment (*DS*, detachment pipetting strength), the area of the bottom surface to be pipetted (*DL*, detachment pipetting length), and trypsin processing time (*DP*, detachment trypsin period) of a passage. *DP* is a contextual parameter that can only be used to perform experiments at fixed values, owing to the specifications of the experimental system. In this case, *DP* is allowed to take different fixed values at three-minute intervals, corresponding to the number of wells in the plate. From the induction step on DDays 1 to 25, I selected two parameters to optimize: the concentration of KnockOut Serum Replacement (KSR) in the medium (*KP*, KSR period), and

the duration of exposure period of the three chemical supplements (*3P*, three supplement period).

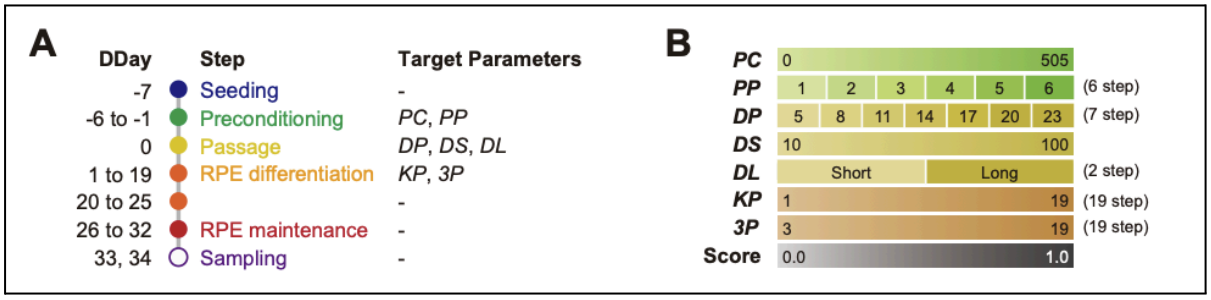


Figure 2.9: Target parameters and corresponding steps in the protocol.
(**A**) Definition of the target parameters and corresponding steps in the protocol: *PC*, preconditioning concentration; *PP*, preconditioning period; *DP*, detachment trypsin period; *DS*, detachment pipetting strength; *DL*, detachment pipetting length; *KP*, KSR concentration reducing period; and *3P*, three chemical (Y, SB, CKI) supplement administration period.
(**B**) Ranges and stepping of the parameters.

Parameter name	Code	Description	Range	Unit	Prptocol step
Preconditioning concentration	*PC*	FGFRi concentration in medium	10–505	nM	Preconditioning
Preconditioning period	*PP*	FGFRi duration in medium	1–6	day	Preconditioning
Detachment trypsin period	*DP*	Trypsin incubation duration at room temperature after incubation at 37°C, 20 min.	5, 8, 11, 14, 17, 20, 23	min	Passage
Detachment pipetting strength	*DS*	Pipetting strength during cell detachment	10–100	mm/sec	Passage
Detachment pipetting length	*DL*	Bottom surface area to be pipetted	short / long	N/A	Passage
KSR period	*KP*	KSR concentration and duration in medium: KSR concentration is decreased linearly every day so that KSR becomes 10% on DDday of KP value	1–19	day	RPE differentiation
Three supplements period	*3P*	Three chemical supplements duration	3–19	day	RPE differentiation

Table 2.1. Definition of optimized parameters.
Parameter names, parameter name codes, description, parameter ranges, parameter units, correspondence between experimental procedure and parameters used (related to **Figures 2.4, Figure 2.9**).

2.4.3. Optimization of the protocol

To improve the optimization performance, 48 conditions (eight plates × six wells, as shown in **Figure 2.7A**) were executed in parallel in each batch. The 48 conditions were selected from the search space using the Bayesian optimization module to maximize the acquisition function calculated from the past experimental data. In general, solving a high-dimensional, expensive black-box optimization problem such as the present one with a limited number of rounds is challenging. In our case, some 200 million possible parameter combinations existed in the search space, and the point where the pigmented score was optimal in three rounds (144 queries) had to be determined, because one experiment round took 40–45 days. In recent studies, BBO has shown excellent performance in real-world black-box optimization problems [6,23,24]. I integrated an experimental design module based on BBO to effectively search for the optimal experimental parameters that maximize the pigmentation scores in the search space defined in **Figure 2.9B**.

The Bayesian optimization module generates queries using two components: the Model updater, which updates the surrogate model that captures the relationship between

parameters and the scores using Bayesian inference (**Figure 2.23**); the Query generator, which generates the next experimental parameters X_{next} using an acquisition function and a policy function (**Figure 2.10A**, **Figure 2.24**; **Algorithms 1–3**). In the Query generator, the acquisition function estimates the expected progress toward the optimal experimental parameter at a given experimental parameter (**Figure 2.10B**). Then, using the acquisition function, the policy function generates the next 48 experimental parameters X_{next} considering the structural context of trypsin processing time x_{DP} (**Figure 2.10C**).

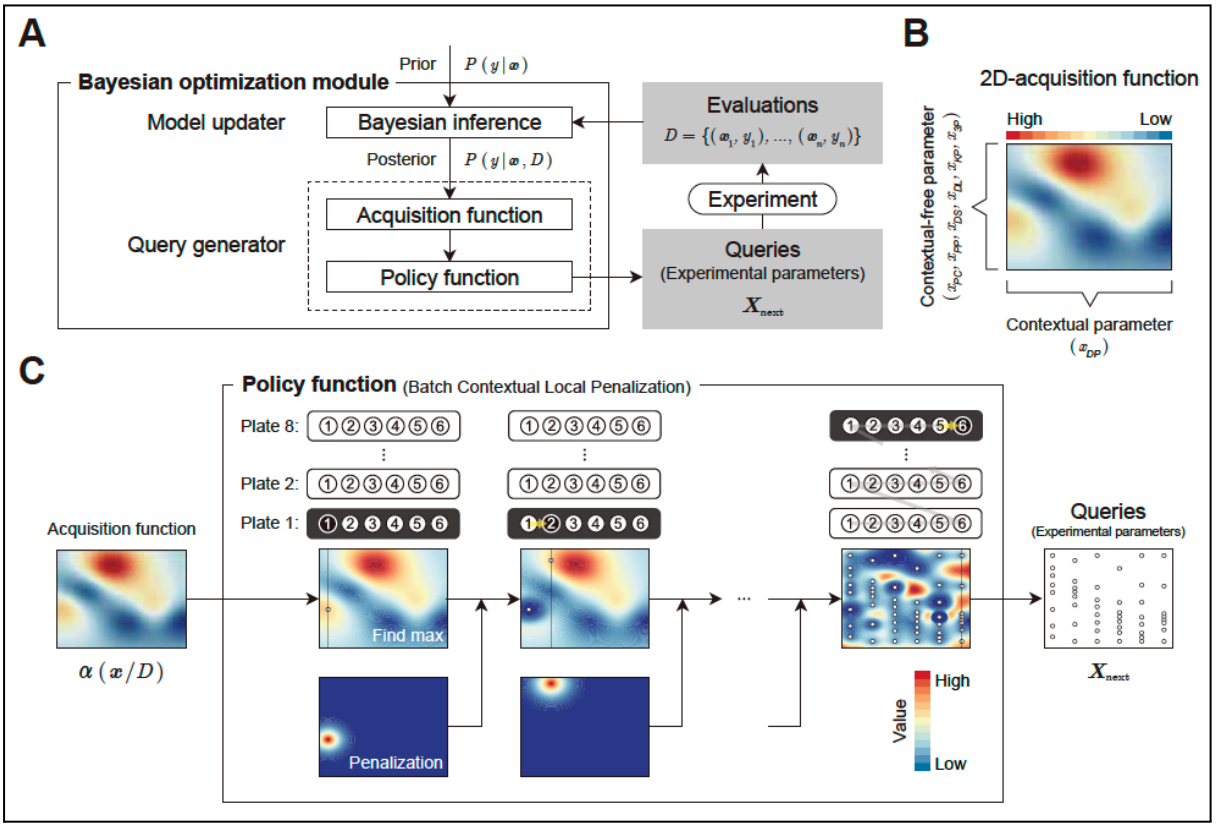


Figure 2.10: Optimization module.

(**A**) The Bayesian optimization module consists of two components: a Model updater and a Query generator. The Model updater updates the Gaussian process posterior on the experiment using all available data $D = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{n}$, where x indicates experimental parameter, and y indicates corresponding evaluation score. The Query generator calculates the acquisition function $\alpha(\boldsymbol{x}; D)$ for an experiment parameter $\boldsymbol{x}$ with the posterior distribution $P(y|\boldsymbol{x}, D)$, and generates the experiment parameter set X_{next} for the next 48 points using the policy function with $\alpha(\boldsymbol{x}; D)$.

(**B and C**) Test of the query generation process using a two-dimensional toy acquisition function.

2.4.4.Robotic optimization drastically improved the pigmentation score

Testing optimization in simulation

A preliminary performance validation of the Bayesian optimization module was conducted by optimizing a seven-dimensional toy testing function constructed on domain

knowledge **(Figure 2.11). Figure 2.12A** shows a comparison between the performance of batch Bayesian optimization and that of random search when the scale of the observation noise on the testing function was changed. The testing function with different noise (SD = 0.000, 0.064, 0.400) was optimized using batch Bayesian optimization with batch contextual local penalization (BCLP) and uniform random sampling, using 8 plates × 6 wells = 48 queries per round. Eighteen independent experiments were performed in each condition. When the observation noise was not present (SD = 0.000) or was sufficiently small (SD = 0.064), as in the proposed system shown in **Figure 2.7C**, BCLP shows better convergence performance compared to the random search. When the scale of the observation noise was relatively large (SD = 0.400), the performance of BCLP was comparable to that of the random search case. **Figure 2.12B** compares the performances of BCLP when the batch size was changed. For the benchmark function (noise SD = 0.064) , optimization was performed using a different number of plates (Np) per round. In each condition, I performed n = 18 independent experiments and showed that the convergence performance improved as the number of plates (the batch size) increased.

Crucially, the simulation results confirmed that the proposed optimization strategy could effectively improve the protocol performance within a feasible number of rounds. Given the high retry cost of the actual culture, this quantitative estimation provided the necessary justification to proceed with the physical implementation of the robotic search.

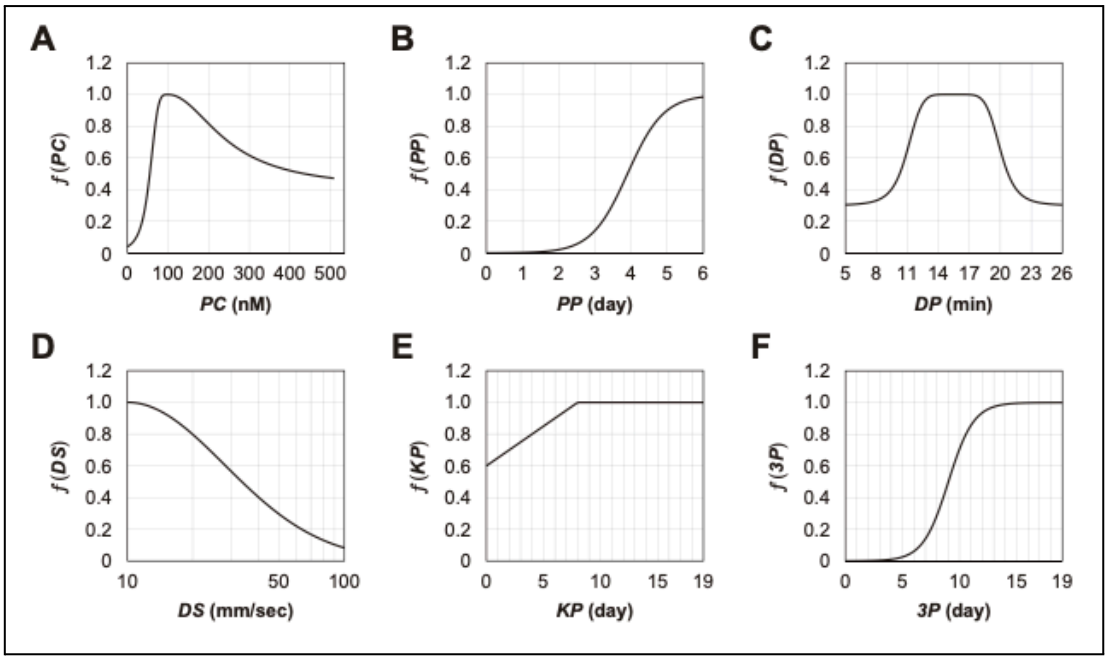


Figure 2.11: Toy testing function based on domain knowledge.
An approximate prediction of how much the value of each variable affects the score based on domain knowledge before performing a series of optimizations. The toy testing function was simply multiplied by the predicted function for each of the variables below, $f(x_{PC}, x_{PP}, x_{DS}, x_{DL}, x_{DP}, x_{KP}, x_{3P}) = f_{PC}(x_{PC}) f_{PP}(x_{PP}) f_{DS}(x_{DS}) f_{DL}(x_{DL}) f_{DP}(x_{DP}) f_{KP}(x_{KP}) f_{3P}(x_{3P})$
(**A**) Prediction of preconditioning (FGFRi) concentration (*PC*) response. It increased from 0 nM, reached a maximum at 100 nM, decreased after 100 nM, and approached 0.5.
(**B**) Prediction of the preconditioning period (*PP*) response: on Day 1 it was 0, it increased monotonically thereafter, and reached a maximum value on Day 6.

(**C**) Prediction of the detachment trypsin period (*DP*) response: an optimal value between 13 and 20 min was expected.
(**D**) Prediction of the detachment pipetting strength (*DS*) response: it achieved a maximum value at 10 mm/s and then decreased monotonically. Although an important parameter, the detachment pipetting length (*DL*) is always assumed to be 1 because of the difficulty of prediction.
(**E**) Prediction of the KSR period (*KP*) response: it was expected to increase monotonically until Day 8, and to be stationary thereafter.
(**F**) Prediction of the three-supplement period (*3P*) response: it was expected to take a non-zero value and increase monotonically from around Day 3 onwards.

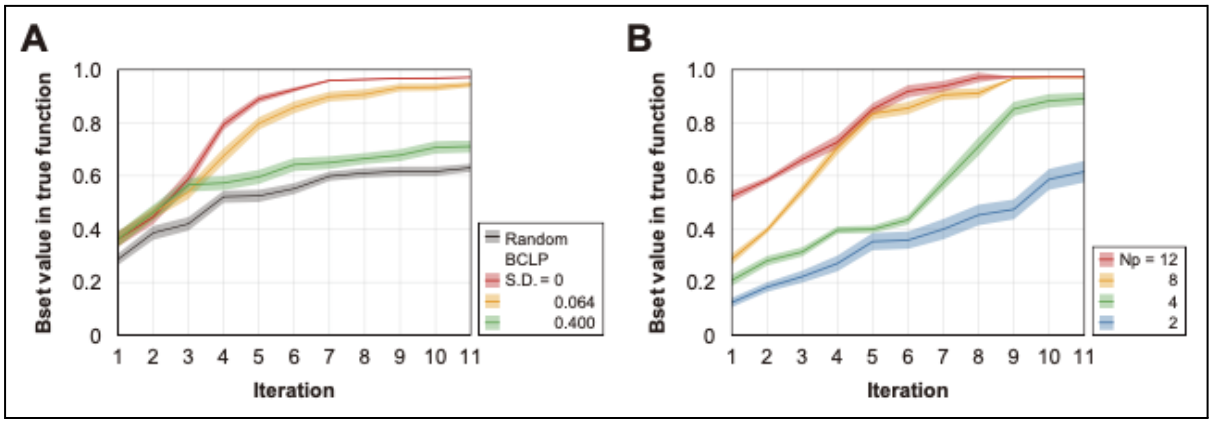


Figure 2.12: Preliminary testing of the Bayesian optimization under different conditions.
The vertical axis shows the value, and the horizontal axis shows the number of circles. The blue horizontal line represents the ground truth (optimal value). In each series, three experiments were conducted independently, and the score of the query with the highest evaluation score in the true function among the queries that had appeared in a certain round was plotted. The error bars represent the standard error (SEM) in each round.
(**A**) The red series shows the results of batch Bayesian optimization (BCLP) with no observation noise. The yellow series shows the results of BCLP with a Gaussian noise SD = 0.064. The green series shows the BCLP results with a high Gaussian noise SD = 0.4. The black series represents the results of random sampling. Compared to random sampling, Bayesian optimization improved the convergence performance and converged to the optimal solution when the observation noise was sufficiently small. Black, random; green, BCLP (SD = 0.4); orange, BCLP (SD = 0.064); red, BCLP (SD = 0). The shaded area represents the SEM image.
(**B**) Comparison of BCLP performance when the batch size was changed. For the benchmark function (noise SD = 0.064), optimization was performed using a different number of plates (Np) per round. The blue series represents Np = 2, the green series represents Np = 4, the orange series represents Np = 8, and the red series represents Np = 16. As the number of plates (the batch size) increased, batch Bayesian optimization tended to converge to an optimal solution with fewer rounds. Blue, Np = 2; green, Np = 4; orange, Np = 8; red, Np = 16. The shaded area represents the SEM image.

Robotic search on the target protocol

In this study, three successive experiments were conducted to optimize the target protocol. In each round, 48 conditions were generated using the Bayesian optimization module and translated into LabDroid operating programs. The robot performed 40 days of iPSC-RPE induction culture under each condition, and I obtained the rate of pigmented cells in the dish as an evaluation score (pigmentation score) for each condition. In accordance with the experimental design, I incorporated the two highest-scoring conditions from the previous experiment (**Figure 2.7C**) as control conditions, performed differentiation-inducing cultures with the LabDroid, and validated the area of the colored cells. In round 1, although one condition was found to be experimentally deficient, the other 47 conditions were validated. The highest score was 0.86 (**Figure 2.13A**, **Figure 2.14**), yielding five conditions that exceeded the mean value (0.39) for all wells in the baseline experiment (**Figure 2.7C**). In round 2, 46 conditions were generated, and the two highest-scoring conditions in round 1 were incorporated as control conditions. The highest score was 0.83 (**Figure 2.13B**, **Figure 2.14**). In round 3, 48 experiments were conducted, yielding an improved highest score of 0.91. I obtained 26 other conditions that were better than the highest in round 2 (**Figure 2.13C**, **Figure 2.14**). A visualization diagram of a two-dimensional partial least squares regression (PLS) clearly revealed that the overall experimental parameters tended to converge in a higher pigmented score direction from rounds 1 to 3 (**Figure 2.13D**). A full visualization of the experimental results using a parallel coordinate plot (PCP) is shown in **Figure 2.14B**.

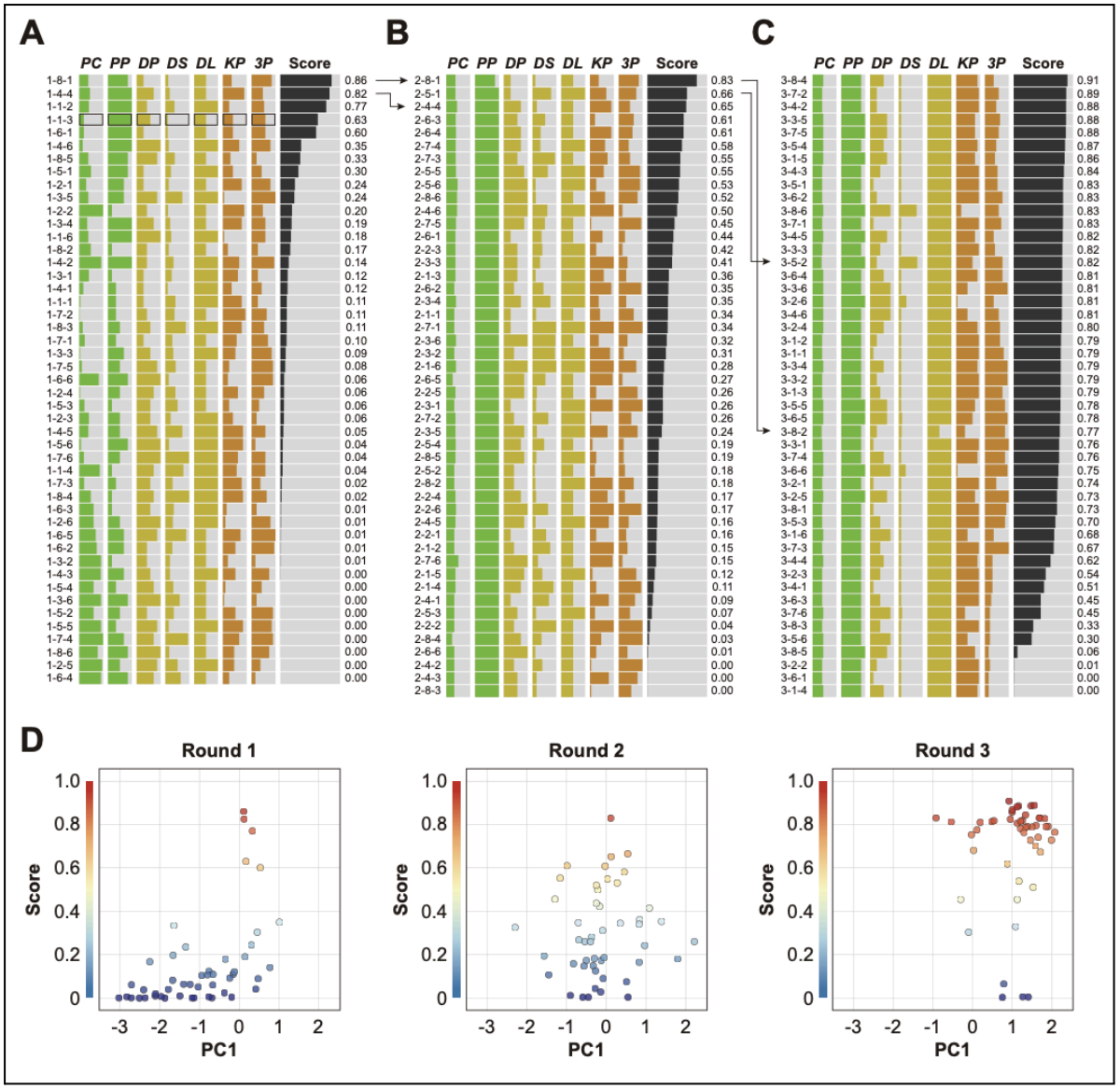


Figure 2.13: Robotic search for optimal parameters in iPSC-RPE differentiation.
(**A–C**) Parameter candidates sorted in order of the pigmentation score in optimization rounds 1 (**A**), 2 (**B**), and 3 (**C**). The ID label on the left represents 'Round No. - Plate No. - Well No.'. For example, “1-2-3” means “(Round) 1-(Plate) 2-(Well) 3.” The parameter values and resulting pigmentation scores are plotted as horizontal bars. The parameter candidate with black frames (1-1-3) in (**A**) is the standard condition. Arrows indicate the control experiments; the top two conditions in round 1 were included in round 2, and the top two conditions in round 2 were implemented in round 3.
(**D**) Visualization of the parameter set and the pigmentation score distributions using partial least squares regression (PLS) in each round. The horizontal axis PC1 shows the values of the parameter candidates that are projected onto the first component of the PLS. The vertical axis shows the pigmentation score for each candidate parameter. As the rounds progressed, the overall score tended to converge in a higher direction.

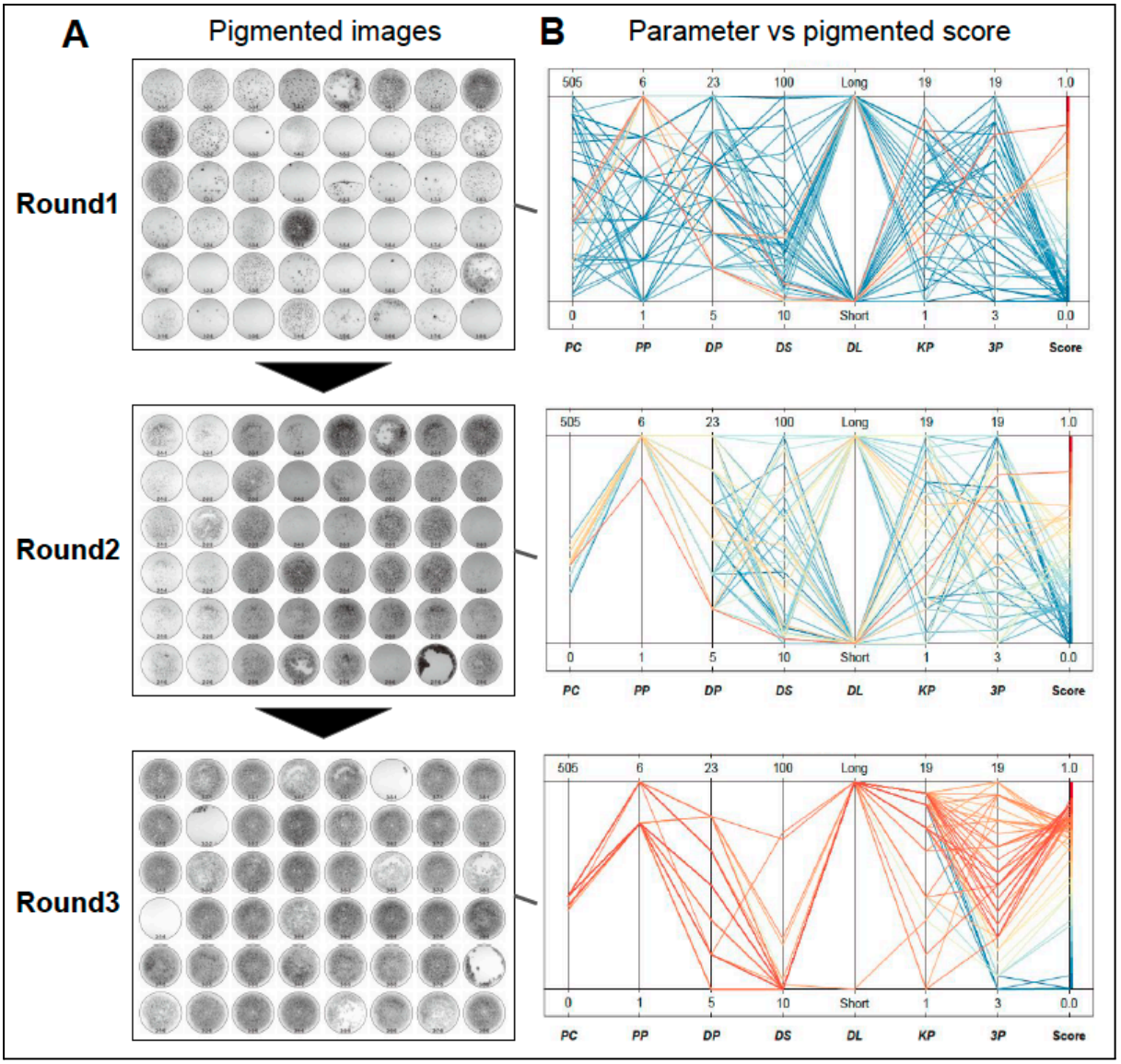


Figure 2.14: Acquired pigmented images of the round 1-3 and Parallel coordinate plot (PCP) of the robotic search experimental results.

(**A**) Images acquired on Day 34 of the round 1-3 experiment: images of the bottom of the well with cultured cells, cropped to the size of the well. These 8-bit images were adjusted to a minimum and maximum contrast value of 100 and 150, respectively. ID labeling on the bottom indicates 'Round No. - Plate No. - Well No.'. (**B**) The experimental results were visualized in rounds 1 to 3 using a parallel coordinate plot (PCP) and represented in 8-dimensional space (7-dimensional parameters + pigmentation scores) as a colored line with vertices on the parallel axes; the position of the vertex on the i-th axis corresponds to the i-th coordinate of the parameter. The color of the line represents the pigmented score: blue lines represent lower pigmentation scores, and the closer to red, the higher the pigmented score.

Validation experiments

To determine whether the optimized conditions were statistically improved over the pre-optimized conditions, an additional multi-well validation experiment was conducted after round 3 using the top five conditions in round 3 and the pre-optimized conditions. The validation values, ordered by place, were 0.71 ± 0.06, 0.72 ± 0.03, 0.76 ± 0.02, 0.79 ± 0.02, and 0.81 ± 0.02 (mean ± SEM, n = 3 each). All scores after optimization were statistically significantly higher than the pre-optimization scores (0.43 ± 0.02; mean ± SEM, n = 3) (**Figure 2.15A, B**). In this study, I succeeded in replacing part of the process of iPS cell differentiation into RPE cells for transplantation using robots, and demonstrated an effective optimization method (**Figure 2.2**). However, it was unclear whether robot-manufactured RPE cells would have the characteristics required for transplantation. Therefore, I purified the cells of the validation round, prepared them for transplantation, and performed a biological quality evaluation. The analyzed iPSC-RPE cells expressed *BEST1*, *RPE65*, and *CRALBP* (**Figure 2.15C**), which are characteristic marker genes of RPE cells. In addition, I observed secretion of VEGF and PEDF into the culture medium, a characteristic of RPE cells (**Figure 2.15D, E**). The expression of tight junction-associated factor ZO-1 was examined using immunohistochemistry, and a ZO-1-derived fluorescence signal was observed in microphthalmia-associated transcription factor (MITF)-positive cells, which play a central role in RPE cell function (**Figure 2.15F**). These results indicated that the robot-manufactured iPSC-RPE cells had the characteristics of RPE cells, and fulfilled the criteria for use in regenerative medicine research using the type of analysis measured in a previous clinical study [21].

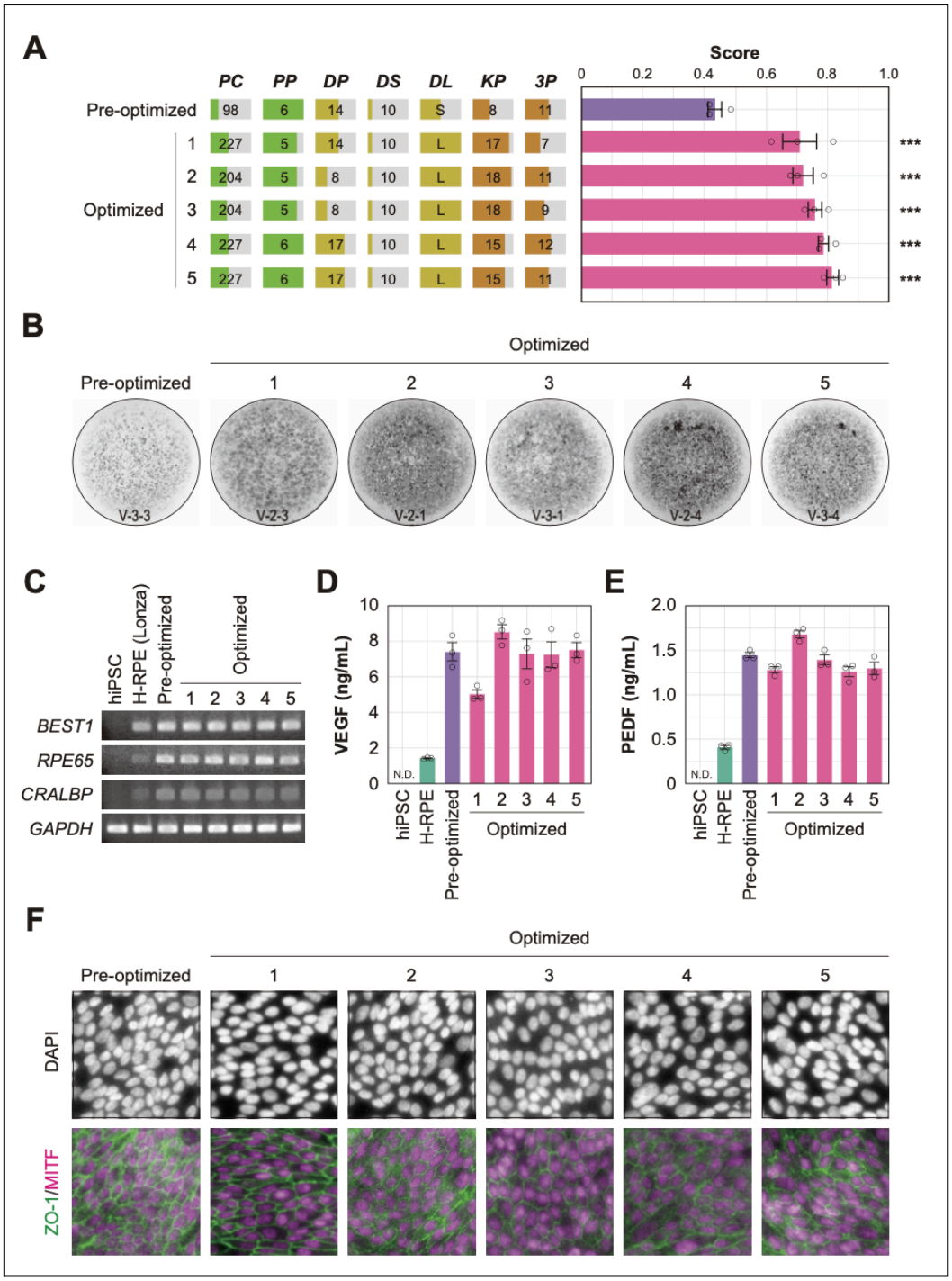


Figure 2.15: Quality evaluation of robot-induced RPE cells.

(**A**) The pigmentation score evaluation of the pre-optimized conditions (n = 3) and the top five conditions from round 3. Error bars represent the standard error of the mean (SEM). The numbers 1–5 in the optimized group represent the first to fifth place conditions for round 3 (**Figure 2.13C**). Circles represent an individual score, bars represent the mean score, and error bars represent the SEM. Statistical significance was examined using two-way ANOVA and SNK *post-hoc* tests. $P < 0.05$ was considered significant. ***$P < 0.001$ versus pre-optimized. In all other combinations no statistical significance was detected.

(**B**) Representative pigmented images of the pre-optimized and five optimized iPSC-RPE cells. Images acquired on DDay 34. ID labeling on the bottom reads 'V (validation) - Plate No. - Well No.'.

(**C–F**) Cell biological validation of the robot-induced RPE cells. After DDay 34, cells were purified, stocked, initiated, maintained for four weeks, and analyzed.
(**C**) Representative marker gene expression in RPE cells by RT-PCR. iPSC, undifferentiated iPSC; H-RPE (Lonza), Clonetics H-RPE (Lot #493461, Lonza, USA); pre-optimized and optimized LabDroid-induced RPE.
(**D–E**) Quantification of representative secreted proteins from iPSC-RPE cells using ELISA. The supernatants were collected and the amount of VEGF (**D**) and PEDF (**E**) in the culture medium was analyzed 24 h after medium exchange (n = 3 wells each). Circles represent individual scores, bars represent the mean score, and error bars represent SEM. n.d. = not detected.
(**F**) Co-staining of ZO-1 (green) and MITF (magenta) using immunohistochemistry. Nuclei were stained with DAPI.

2.5 Discussion

In this study, I proposed a robotic search system to autonomously search for optimal cell culture conditions, bringing together experimental robotics and BBO. Our robotic search system autonomously discovered the optimal combination of seven parameters comprising the iPSC-RPE induction process (target process) required to increase the number of pigmented cells (pigmentation score). Our approach can be applied to cell culture protocols other than iPSC-RPE induction; however, it may not be optimal even when implemented with a completely identical hardware-software setup. Below, I discuss some considerations and potential limitations for tailoring the components of our robotic search system (robots, parameters, and evaluation scores) to other targets.

Robots: the requirements depend on the nature of the target process. The search parameters must be changeable (flexibility), non-search parameters must remain stable or change only within the range of the specifications (reliability), and the operation must be sufficiently repeatable (accuracy). In addition, the storage capacity for CO_2 incubators and refrigerators needs to be set in accordance with the number of cell plates that are to be cultured concurrently. For target processes that require long-term culture (i.e., processes that have high retry costs) such as cell differentiation induction, the robots and peripheral equipment need to have low error rates. In target processes that have low retry costs, a lower priority on low error rates is required. I chose LabDroid for this research, as it meets these requirements and has good future operational extensibility.

Parameters: the number and range of searchable parameters is constrained by the number of experiments that can be performed. The more parameters to be searched, the greater the number of experiments required for sufficient optimization. The available experimental resources (number of iterations or parallel cultures) should be considered in advance for appropriate parameter optimization. Here, I limited the scope of our search to just seven parameters (**Table 2.1**). However, a myriad of potential parameter candidates, including other chemicals, culture media, and order of manipulations, can be considered. During parameter selection, I referred to previous cell culture studies and expert opinions, as well as preliminary simulations, to confirm that optimization was sufficiently feasible with our resources (**Figure 2.12**). The search ranges for the seven parameters were carefully selected for our target process; different appropriate search ranges should be selected in case of other target processes, including the induction of differentiation into other types of tissues.

Evaluation scores: since the optimization is performed on the evaluation scores, designing the evaluation function is critical. Here, I used the pigmentation score as the evaluation score because of the following reasons: when preparing iPSC-RPE cells for transplantation in clinical research, a clinical team evaluates the rate of pigmented cells, gene expression, and secretory substances in cells subjected to differentiation induction followed by purification (**Figure 2.3**). This quality assessment is not based on a total score, and only those cells that satisfy all the criteria in all items are suitable for transplantation [19,21,22]. Because cell pigmentation is one of the criteria for the assessment, cell pigmentation alone is not sufficient to determine cell quality, but can be a requirement. It should be noted that the pigmentation score does not reflect the degree of pigmentation in individual cells, but indicates the number of cells in the dish whose pigmentation is above the threshold. Since pigmented cells and non-pigmented cells are mixed in the dishes at the end of the induction (i.e., before purification), single-cell omics analysis is needed to accurately evaluate the quality of individual cells. For example, in stem cells, a value (stemness index) has been proposed to evaluate stemness from single-cell mRNA-seq information [25]. I believe that if a similar index for iPSC-RPE cells indicating cell quality from transcriptome data is established, this could replace the pigmentation score that I used, and would make the process I have developed even more ideal.

2.6 Materials and Methods

Guidelines

All experiments that involved the use of human-derived samples were reviewed and approved by the institutional review board of the Institutional Committee of RIKEN Kobe Branch.

Reagents

hiPSC maintenance medium: 80% StemFit Basal Solution A and 20% StemFit iPS Expansion Solution B (#AK02N, Ajinomoto Co., Inc., Japan).

RPE differentiation medium (20% KSR): 0.10 mM MEM non-essential amino acids solution (NEAA) (#11140050, Thermo Fisher Scientific Inc., MA, USA), 1.0 mM sodium pyruvate (#S8636, Merck & Co., Inc., NJ, USA), 19% knockOut serum replacement (KSR) (#10828028, Thermo Fisher Scientific Inc., MA, USA), 0.0007% 2-mercaptoethanol (#139-06861, FUJIFILM Wako Pure Chemical Corporation, Japan), 78 U/mL benzylpenicillin sodium, and 78 μg/mL streptomycin sulfate (#15140122, Thermo Fisher Scientific Inc., MA, USA). All diluted in GMEM (#11710035, Thermo Fisher Scientific Inc., MA, USA).

RPE differentiation medium (15% KSR): 0.10 mM MEM NEAA (#11140050, Thermo Fisher Scientific Inc., MA, USA), 0.99 mM sodium pyruvate (#S8636, Merck & Co., Inc., NJ, USA), 15% KSR (#10828028, Thermo Fisher Scientific Inc., MA, USA), 0.0007% 2-mercaptoethanol (#139-06861, FUJIFILM Wako Pure Chemical Corporation, Japan), 82 U/mL benzylpenicillin sodium, and 82 μg/mL streptomycin sulfate (#15140122, Thermo Fisher Scientific Inc., MA, USA). All diluted in GMEM (#11710035, Thermo Fisher Scientific Inc., MA, USA).

RPE differentiation medium (10% KSR): 0.094 mM MEM NEAA (#11140050, Thermo Fisher Scientific Inc., MA, USA), 0.94 mM sodium pyruvate (#S8636, Merck & Co., Inc., NJ, USA), 10% KSR (#10828028, Thermo Fisher Scientific Inc., MA, USA), 0.0007% 2-mercaptoethanol (#139-06861, FUJIFILM Wako Pure Chemical Corporation, Japan), 85

U/mL benzylpenicillin sodium, and 85 μg/mL streptomycin sulfate (#15140122, Thermo Fisher Scientific Inc., MA, USA). All diluted in GMEM (#11710035, Thermo Fisher Scientific Inc., MA, USA).

RPE maintenance medium: 29% Nutrient Mixture F-12 (#N6658, Merck & Co., Inc., NJ, USA), 1.9 mM L-glutamine (#G7513, Merck & Co., Inc., NJ, USA), 1.9% B-27 supplement, serum free (#17504044, Thermo Fisher Scientific Inc., MA, USA), 96 U/mL benzylpenicillin sodium, and 96 μg/mL streptomycin sulfate (#15140122, Thermo Fisher Scientific Inc., MA, USA). All diluted in DMEM (Low glucose) (#D6046, Merck & Co., Inc., NJ, USA).

FGF receptor inhibitor (FGFRi) stock: PD 173074 (#P2499-5MG, Merck & Co., Inc., NJ, USA) diluted in DMSO (#D2650-5X5ML, Merck & Co., Inc., NJ, USA).

Rho-kinase inhibitor (Y) stock (8–10 mM): CultureSure Y-27632 (#036-24023, FUJIFILM Wako Pure Chemical Corporation, Japan) diluted in distilled water (Otsuka Pharmaceutical Factory, Japan) to a final 10 μM concentration when added to the cell culture medium.

TGF-β/Activin/Nodal signal inhibitor (SB) stock (4–5 mM): SB 431542 hydrate (#S4317-5MG, Merck & Co., Inc., NJ, USA) diluted in DMSO (#D2650-5X5ML, Merck & Co., Inc., NJ, USA) to a final 5 μM concentration when added to the cell culture medium.

Wnt signal inhibitor (CK) stock (2.4–3 mM): CKI-7 dihydrochloride (#C0742-5MG, Merck & Co., Inc., NJ, USA) diluted in distilled water (Otsuka Pharmaceutical Factory, Japan) to a final 3 μM concentration when added to the cell culture medium.

RPE adhesion medium: DMEM/F12 (D8437, Merck & Co., Inc., NJ, USA), 10% FBS (12007C, Nichirei Corporation, Japan).

RPE washing solution: 98% DMEM/F12 (D8437, Merck & Co., Inc., NJ, USA), 1 mM sodium pyruvate (S8636, Merck & Co., Inc., NJ, USA), 2 mM L-glutamine (G7513, Merck & Co., Inc., NJ, USA).

Labware

For human use: micropipette tip, 2140-05-HR/2149P-05/61849, Thermo Fisher Scientific Inc. (MA, USA); micropipette tip, 30389165, Mettler Toledo (OH, USA); micropipette tip, 737251, Greiner Bio-One International GmbH (Germany); disposable pipette, 356507, Corning Incorporated (NY, USA); disposable pipette, 606160/607160/760160/768160, Greiner Bio-One International GmbH (Germany); filtration, SLGVJ13SL, Merck & Co., Inc. (NJ, USA); filtration, SS-10LZ, Terumo Corporation (Japan); filtration, 431096/430281/431097/430282, Corning Incorporated (NY, USA); 1.5-mL tube, 72.692MS, Sarstedt K.K. (Japan); 15-mL tube, 352096, Corning Incorporated (NY, USA); 50-mL tube, 352070, Corning Incorporated (NY, USA).

For LabDroid use: 6-well plate, 353046, Corning Incorporated (NY, USA); 50 mL tube, MS-58500, Sumitomo Bakelite Co., Ltd. (Japan); micropipette tip, 3511-05-HR/3512-05-HR/94410313/94410713/94052550, Thermo Fisher Scientific Inc. (MA, USA).

LabDroid Maholo booth

LabDroid including peripheral equipment were placed inside a booth made of acrylic walls and a stainless steel frame with three fan-filter-units (**Figure 2.16**). The LabDroid booth included a dual-arm humanoid (Robotic Biology Institute Inc., Japan), a CO_2 incubator (APC-30D, ASTEC Co., Ltd., Japan), micropipettes

(4641110N/4641030N/4641230N/4641210N, Thermo Fisher Scientific Inc., MA, USA), a tube rack (Robotic Biology Institute Inc., Japan), a plate rack (Robotic Biology Institute Inc., Japan), a dry bath (EC-40RA, AS ONE Corporation, Japan), a tip sensor (Robotic Biology Institute Inc., Japan), an aspirator (SP-30, Air Liquide, Italy), a dust bin (EPD3S, Sekisui Techno Moulding Co., Ltd., Japan), and a microscope (EVOS FL Auto 2, Thermo Fisher Scientific Inc., MA, USA).

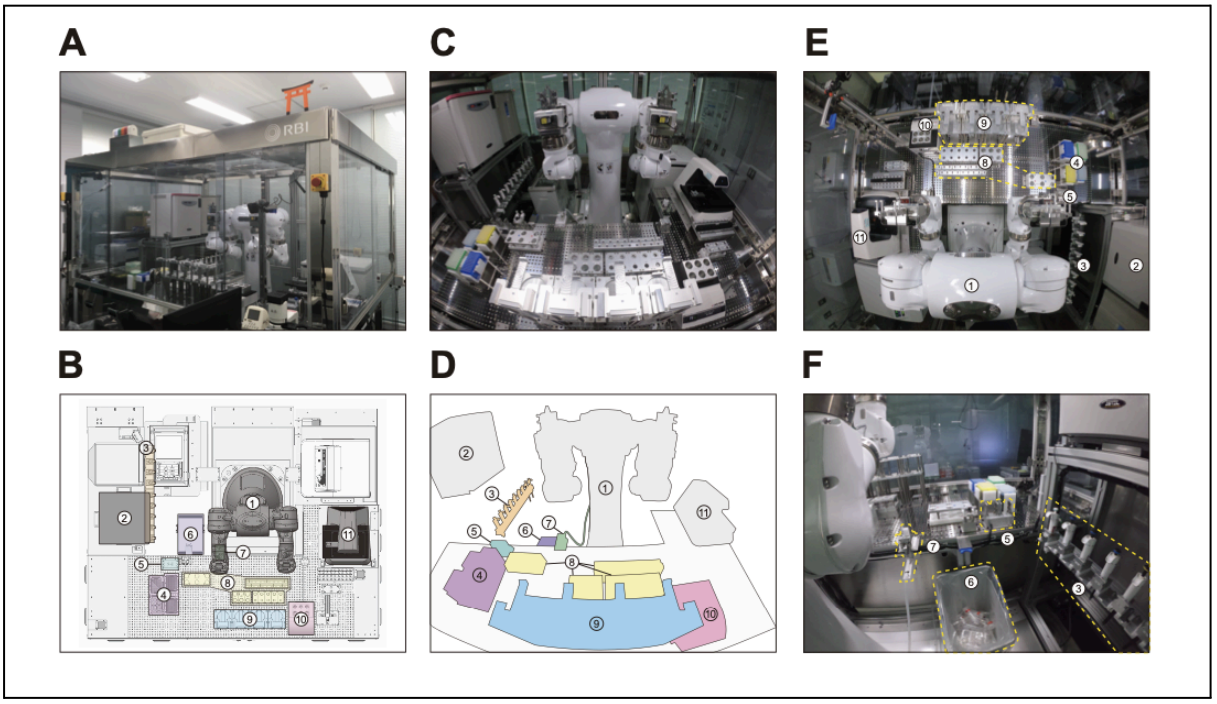


Figure 2.16: LabDroid Maholo booth.
(**A**) Exterior photograph of the LabDroid booth. The LabDroid consists of an acrylic box of W2500 x D2000 x H2200 (mm).
(**B**) Plan view of the LabDroid booth (3D-CAD) and layout of the equipment.
(**C–D**) Front view photograph of the LabDroid booth and schematic drawing of the equipment (these panels are identical to those in **Figure 2.5**).
(**E**) Top view photograph of the LabDroid booth.
(**F**) A back view photograph of the LabDroid booth.
Components: (1) dual-arm humanoid; (2) CO_2 incubator; (3) micropipettes; (4) pipette tips; (5) tip sensor; (6) dustbin; (7) aspirator; (8) tube racks; (9) plate racks; (10) dry bath; and (11) microscope.

---

hiPSC culture — initiation and preparation of cell suspensions (human part)

The hiPSC line 253G1 [26], made from human dermal fibroblasts, was obtained from RIKEN BRC (HPS0002). The hiPSCs were cultured and differentiated using the method previously described [27–29].

On DDay −14, frozen hiPSCs were initiated using the following procedures: first, laminin-coated 6-well plates were prepared. A final concentration of 0.5 μg/cm^2 iMatrix-511 (Matrixome Inc, Japan) diluted in PBS (-) was then added to each well of the four 6-well

plates and incubated for a minimum of 60 min at 37 ˚C and 5% CO_2, after which 0.75 mL/well of hiPSC maintenance medium was added. The supernatant was then removed. Next, 1 mL/well of hiPSC maintenance medium containing Rho-kinase inhibitor (final 10 μM concentration) was added, and the coated plates were incubated at 37 °C and 5% CO_2 until further use.

For hiPSC initiation, frozen vials of hiPSCs stored in liquid nitrogen were thawed in a water bath set at 37 °C, and the cells were subsequently suspended in 5 mL of hiPSC maintenance medium. After centrifugation (160 × *g*, 22 °C, 4 min), the supernatant was removed and an appropriate volume of hiPSC maintenance medium with a final 10 μM Rho-kinase inhibitor concentration was added. After counting the cells with a hemocytometer, the cells were seeded into laminin-coated 6-well plates at 43,300–45,000 cells/1.5 mL medium/well.

On DDay −13, the medium was replaced with hiPSC maintenance medium without Rho-kinase inhibitor. On DDays −12 to −8, the medium was replaced with the same medium composition at 24–72 h intervals. On DDay −7, cells were collected from the plate, and cell suspensions were delivered to the LabDroid booth. The medium was aspirated and 2 mL/well of PBS (-) was gently added and then aspirated for washing, followed by addition of 1 mL of 0.5 x TrypLE Select CTS (#A12859-01, Thermo Fisher Scientific Inc., MA, USA) diluted in 0.5 mM EDTA/PBS (-) and incubated at 37 °C and 5% CO_2 for 10 to 20 min. Then, cells were detached by pipetting and collected into a 50 mL tube, to which 1 mL of hiPSC maintenance medium and 3 mL of PBS (-) were added. After centrifugation (160 × *g*, 22 °C, 4 min), the supernatant was removed, 0.75 mL of hiPSC maintenance medium with 10 μM Rho-kinase inhibitor was added, and the cells were resuspended. The cell suspension was filtered through a 40 μm cell strainer (#352340, Corning Incorporated, USA) with an additional 0.75 mL of hiPSC maintenance medium. After counting the cells with a hemocytometer, the cell suspension was set to 133,400 cells/20 mL with hiPSC maintenance medium containing 10 μM Rho-kinase inhibitor in eight 50 mL tubes. To prepare the cell suspensions, eight 6-well plates coated with laminin were prepared. A final concentration of 0.5 μg/cm2 of iMatrix-511 (Matrixome Inc., Japan) diluted in PBS (-) was added to each well of four 6-well plates and incubated for a minimum of 60 min at 37 °C and 5% CO_2.

iPSC-RPE differentiation (LabDroid part)

On DDay −7, the hiPSC suspension was seeded into eight 6-well plates by coating eight 6-well plates with laminin, and placing eight tubes of the iPSC suspension and labware in the appropriate positions. The task of seeding was initiated, and the robotic operation was performed by LabDroid (**Figure 2.6A**, **Figure 2.17**). After the robotic operation, the eight cell-seeded plates were exported and incubated in a CO_2 incubator outside the LabDroid booth.

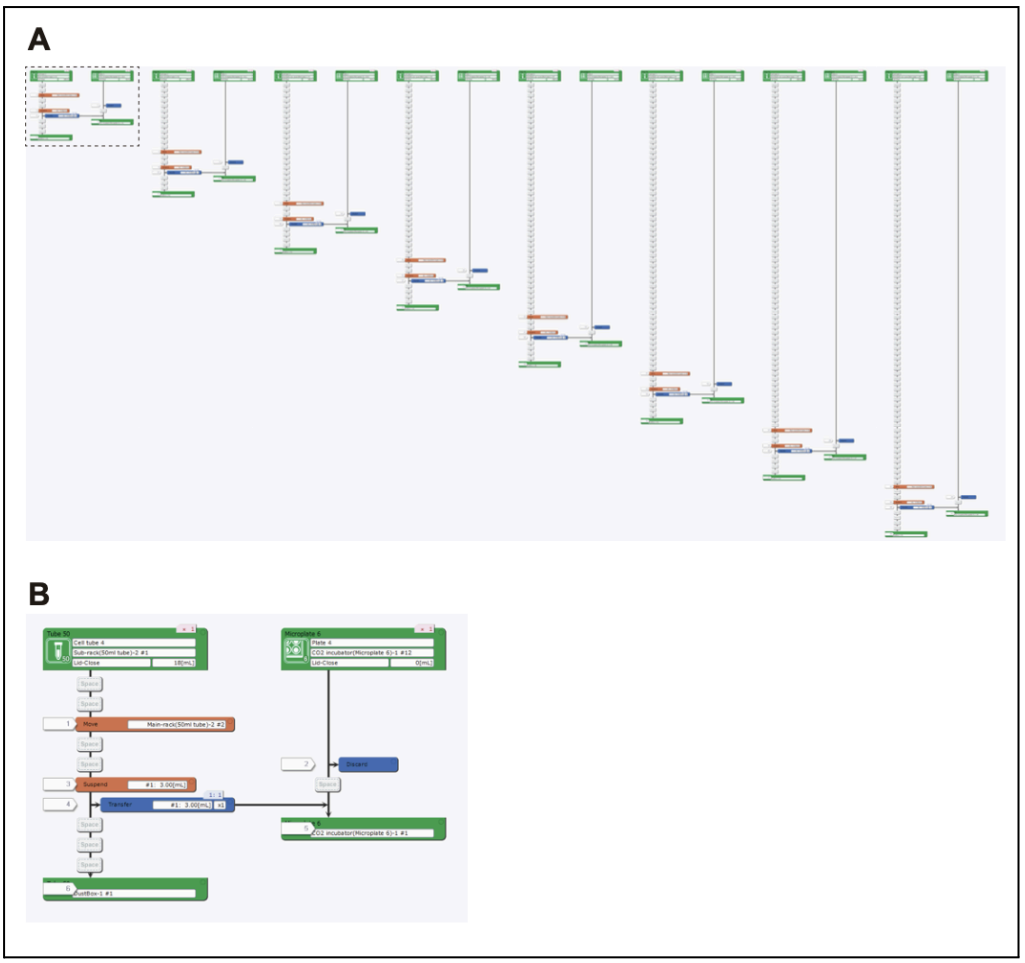


Figure 2.17: Representative LabDroid execution of a seeding experiment (round 3, DDay −7)
(**A**) Entire image. (**B**) Enlarged image of the dotted rectangle from panel **A**.

On DDay −6, the eight seeded plates were imported into the CO_2 incubator of the LabDroid booth. The users prepared eight 50 mL tubes of hiPSC maintenance medium with a final 10 μM Rho-kinase inhibitor concentration and two 50 mL tubes of hiPSC maintenance medium with final 5 μM FGFRi and 10 μM Rho-kinase inhibitor concentrations. The reagents and labware were placed in the appropriate positions. The task of preconditioning was then initiated, and the robotic operation was performed by LabDroid (medium exchange type I; **Figure 2.6B**, **Figure 2.18**).

On DDays −5 to −1, the users prepared eight 50 mL tubes of hiPSC maintenance medium without Rho-kinase inhibitor and two 50 mL tubes of hiPSC maintenance medium with a final 5 μM FGFRi concentration. The reagents and labware were placed in the appropriate positions. The task of preconditioning was initiated, and the robotic operation was performed by LabDroid (medium exchange type I;**Figure 2.6B**, **Figure 2.18**).

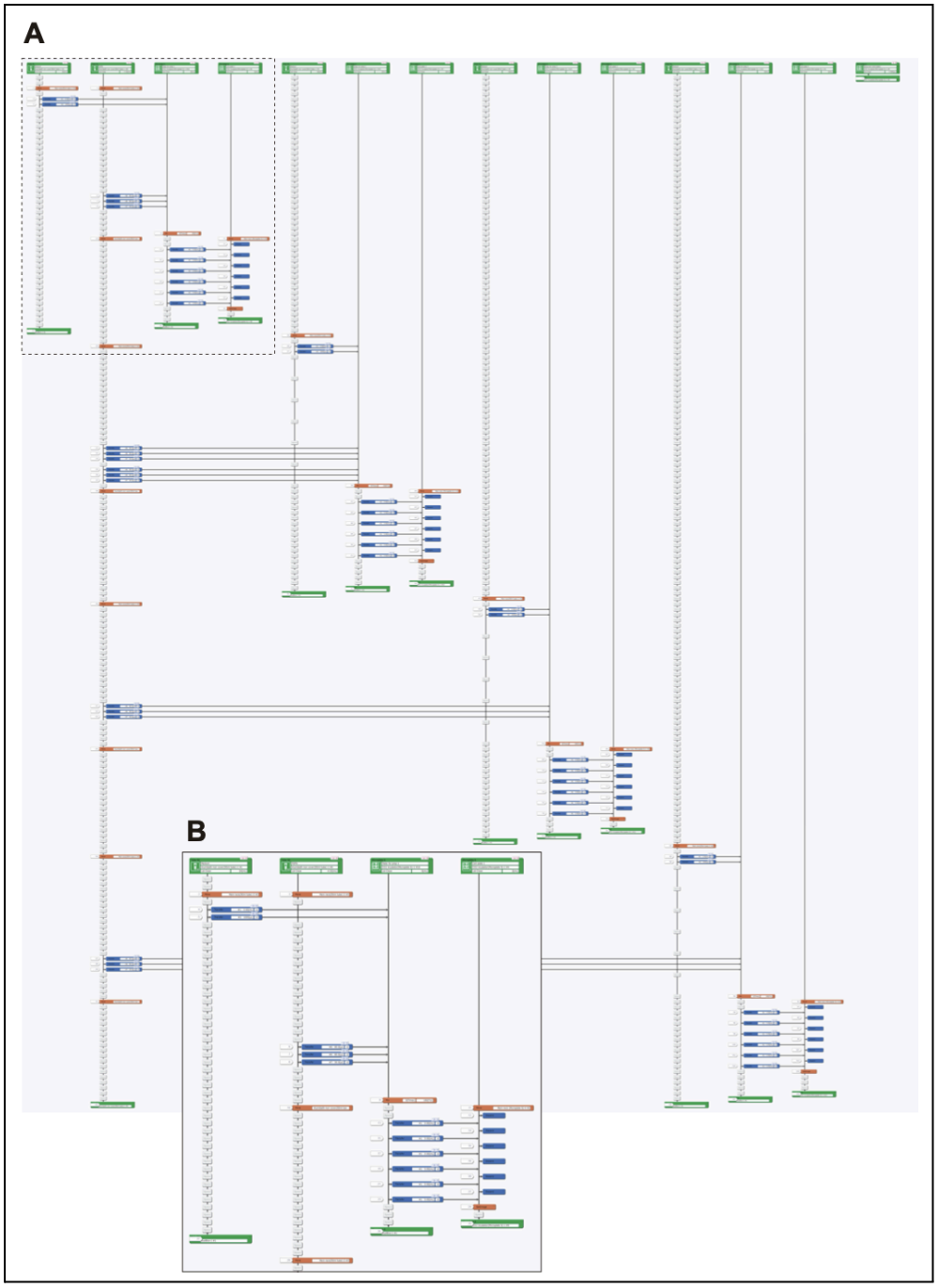


Figure 2.18: Representative LabDroid execution of a preconditioning experiment (round 3, DDay −6, 1st run)
(**A**) Entire image. (**B**) Enlarged image of the dotted rectangle from panel **A**.

On DDay 0, the following procedure was used for the operation of four plates: the users prepared four 6-well plates coated with laminin. A final 0.5 μg/cm^2 concentration of iMatrix-511 (Matrixome Inc., Japan) diluted in PBS (-) was added to each well of the four 6-well plates and then the plates were incubated for a minimum of 60 min at 37 °C and 5% CO_2. The users also prepared two 50 mL tubes of PBS (-), two 50 mL tubes of 0.5 x TrypLE Select CTS (#A12859-01, Thermo Fisher Scientific Inc., MA, USA) diluted in 0.5 mM EDTA/PBS (-), and four plates with RPE differentiation medium (20% KSR) with final 10 μM

Rho-kinase inhibitor/3 μM Wnt signal inhibitor/5 μM TGF-β/Activin/Nodal signal inhibitor (4 mL/well each) concentration. The cell plates, laminin-coated plates, plates with medium, reagents, and labware were placed in the appropriate positions. The task of passage was initiated, and robotic operations were performed by LabDroid (**Figure 2.6D**, **Figure 2.19**). After performing this operation twice (four plates each), the eight cell-passaged plates were exported and incubated in a CO_2 incubator outside the LabDroid booth.

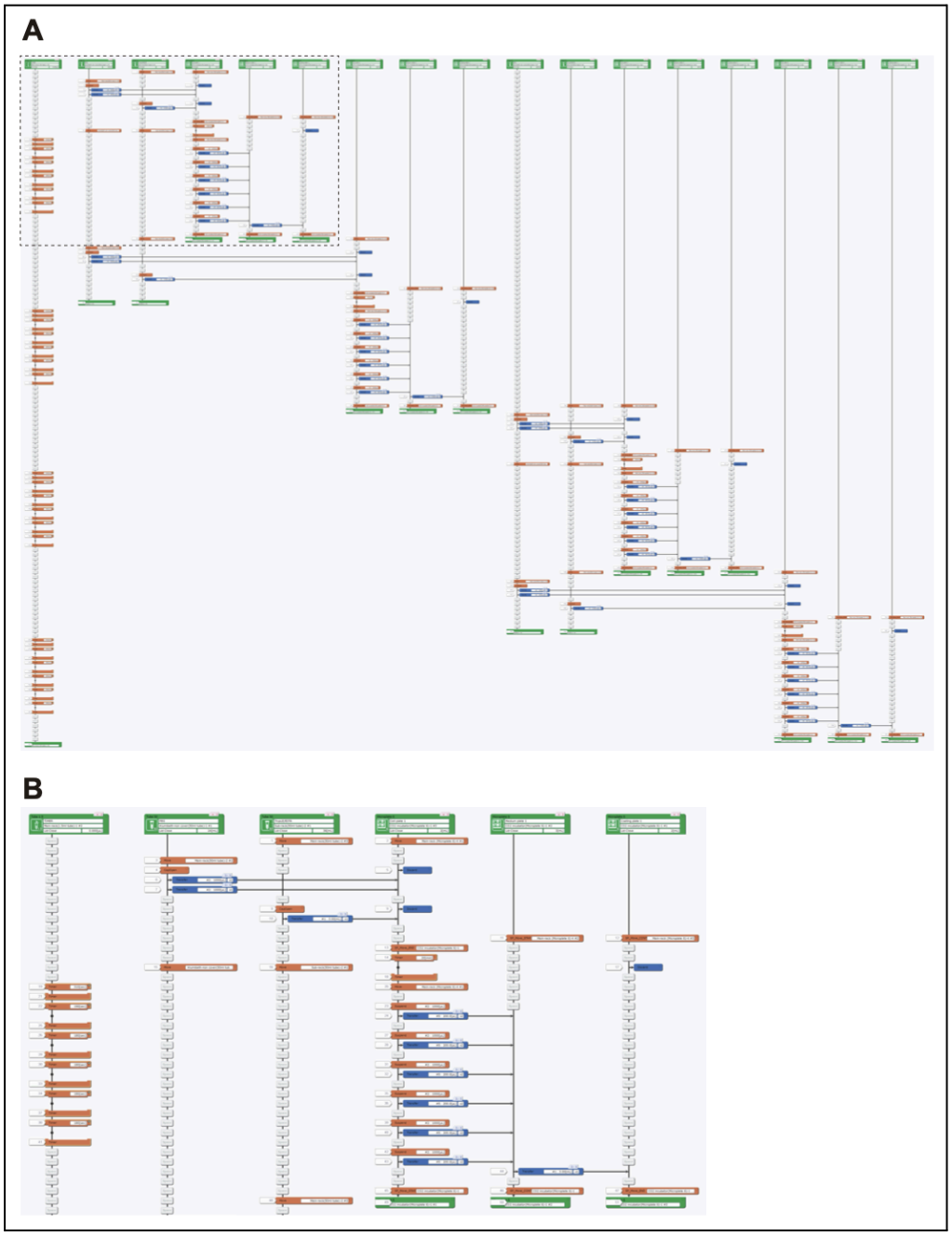


Figure 2.19: Representative LabDroid execution of a passage experiment (round 3, DDay 0, 1st run)
(**A**) Entire image. (**B**) Enlarged image of the dotted rectangle from panel **A**.

On DDay 1, the eight cell-passaged plates were imported into the CO_2 incubator of the LabDroid booth. Users prepared eight 50 mL tubes of RPE differentiation medium (10% KSR), two 50 mL tubes of 100% KSR, one 50 mL tube of 4 mM Rho-kinase inhibitor stock/1.2 mM Wnt signal inhibitor stock, and one 50 mL tube of 4 mM TGF-β/Activin/Nodal signal inhibitor stock. The reagents and labware were placed in the appropriate positions. The task of RPE differentiation was initiated, and the robotic operation was performed by LabDroid (medium exchange type I; **Figure 2.6B**, **Figure 2.20**).

On DDays 2 to 19, the users prepared eight 50 mL tubes of RPE differentiation medium (10% KSR), two 50 mL tubes of 100% KSR, one 50 mL tube of 4 mM Rho-kinase inhibitor stock/1.2 mM Wnt signal inhibitor, and one 50 mL tube of 4 mM TGF-β/Activin/Nodal signal inhibitor. The reagents and labware were placed in the appropriate positions. The task of RPE differentiation was initiated, and the robotic operation was performed by LabDroid (medium exchange type I; **Figure 2.6B**, **Figure 2.20**).

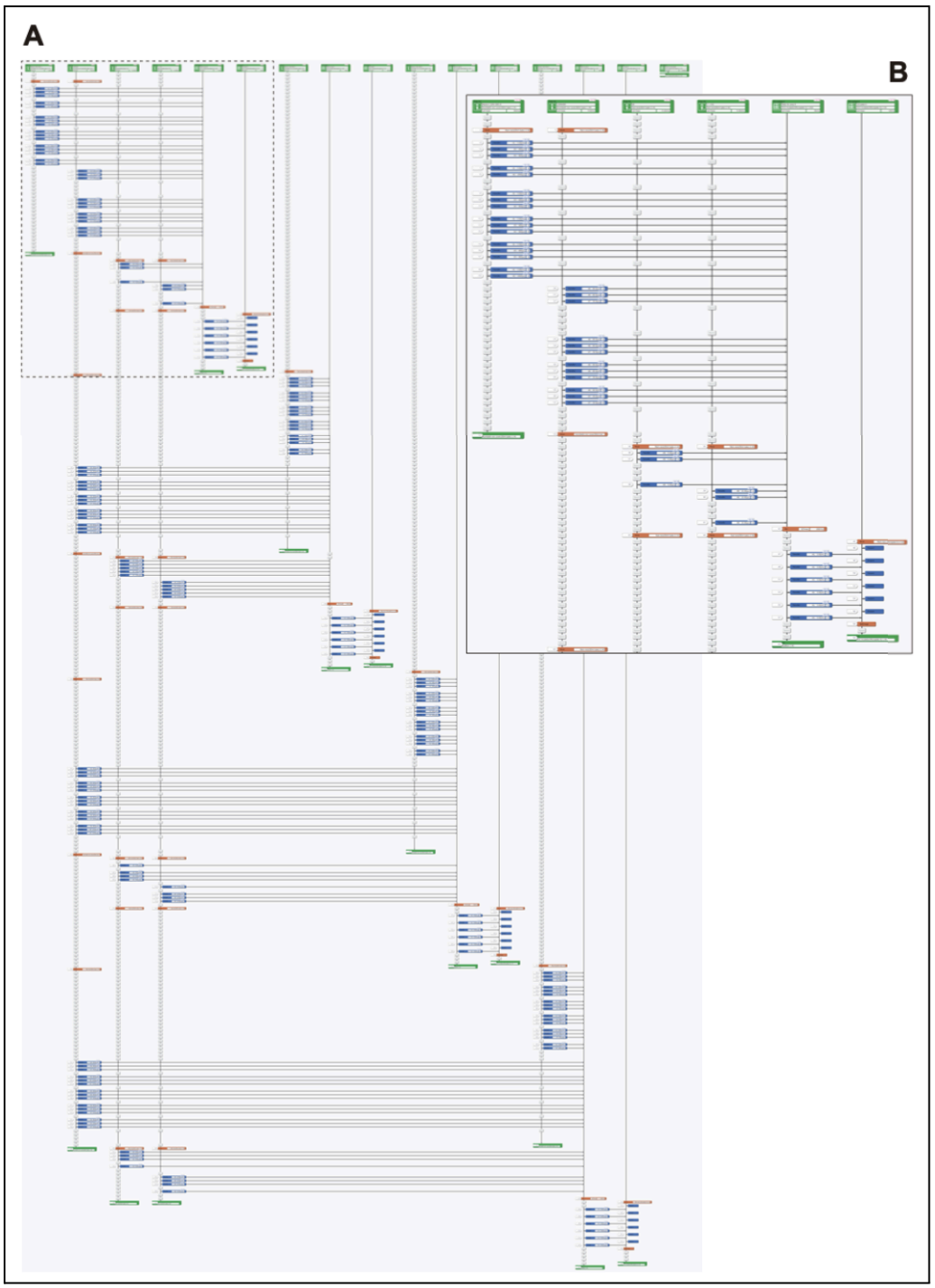


Figure 2.20: Representative LabDroid execution of an RPE differentiation experiment (round 3, DDay 10, 1st run)
(**A**) Entire image. (**B**) Enlarged image of the dotted rectangle from panel **A**.

On DDays 20 to 32, the users prepared eight 50 mL tubes of RPE differentiation medium (10% KSR; DDays 10 to 25) or RPE maintenance medium (DDays 26 to 32). The reagents and labware were placed in the appropriate positions. RPE differentiation and maintenance were initiated and the robotic operations were performed by LabDroid (medium exchange type II; **Figure 2.6C and 2.21**).

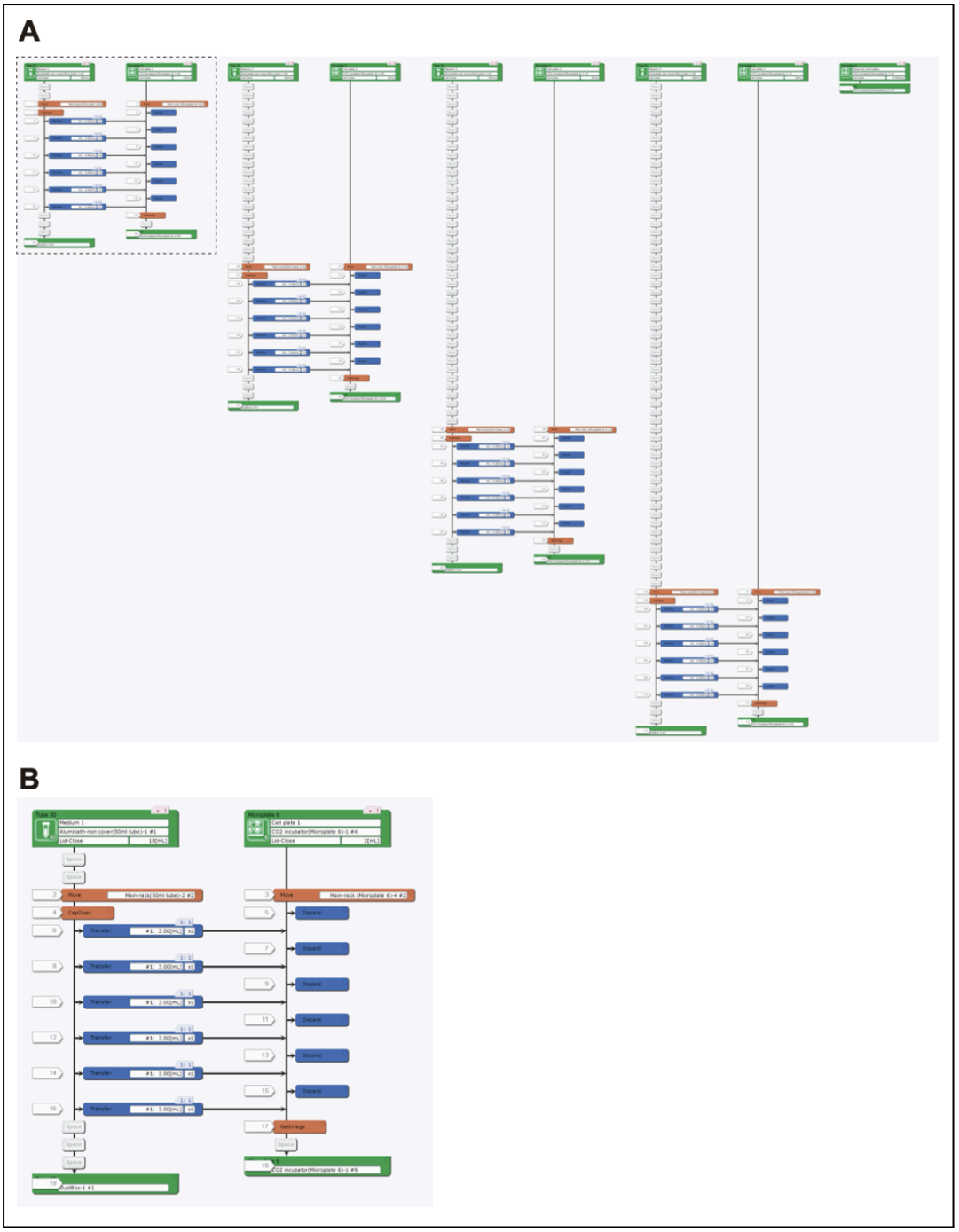


Figure 2.21: Representative LabDroid execution of an RPE maintenance experiment (round 3, DDay 32, 1st run)
(**A**) Entire image. (**B**) Enlarged image of the dotted rectangle from panel **A**.

Scoring — sampling

On DDay 33, the cell plates were exported and the cell culture medium was replaced with fresh RPE maintenance medium. After 24 h (DDay 34), the medium was collected for ELISA analysis. The remaining media were aspirated and 2 mL of PBS (-) were added and then aspirated for washing. After that, photographic images were acquired for the calculation of scoring values.

Scoring — image analysis

Images were acquired using a digital camera (PSG7X MARKII, Canon Inc., Japan): ISO 500; focal length F=9.00, 50 mm; exposure time, 1/1250 sec. The camera was set in the same position throughout all experiments. The acquired images were automatically processed by filtering with Gaussian blur, subtracting the background, binarizing by thresholding with a constant value, and cropping with a constant pixel value. The colored cell area was then calculated (**Figure 2.22**).

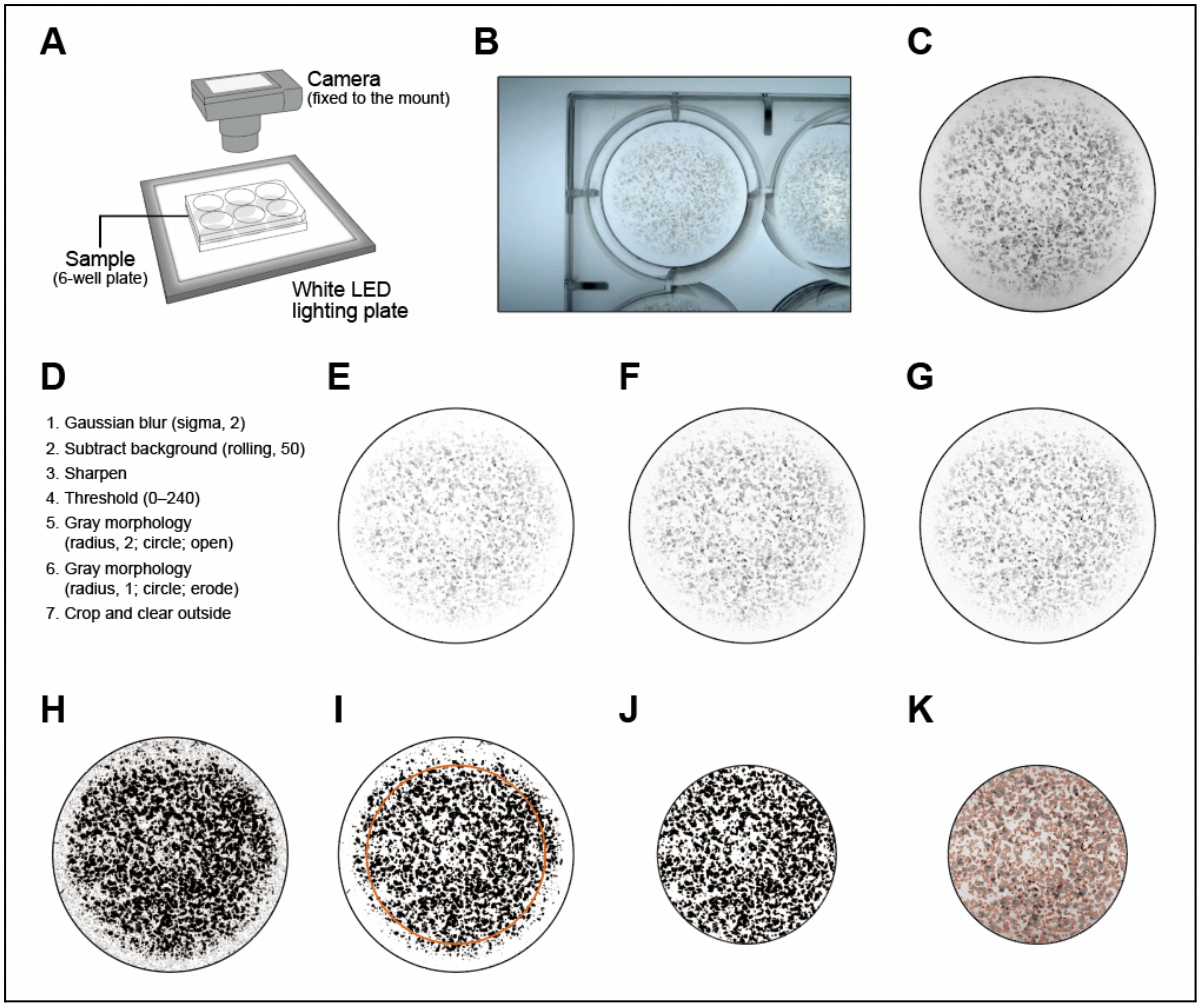


Figure 2.22: Image processing to calculate pigmentation scores

(**A–C**) The 6-well plates were placed on top of a white LED lighting plate and photographed with a camera fixed to the mount (**A**). The raw images acquired (**B**) were cropped in a circle to the size of the bottom of the well (**C**).

(**D**) Workflow of image processing performed using ImageJ/Fiji macro. The background was removed in steps 1–3, binarization in step 4, and noise in steps 5 and 6. Since it is empirically known that differentiation-inducing cells are less likely to grow near the sides of the wells, only the central portion was cropped (step 7) and the scores were subsequently calculated.

(**E–J**) Examples of processed images. The images shown are samples from well 1 of plate 1 in the baseline experiment. Images after Gaussian blur (**E**), background subtraction (**F**), sharpening (**G**), binarization by thresholding (**H**), mathematical morphology processing (**I**),

crop, and clear outside (**J**) processing. The orange circle in panel **I** represents the crop area in step 7. The pigmentation score was calculated as the area of the black region in panel **J**.
(**K**) Merged image of the image before processing (corresponding to panel **C**) and the area determined to be pigmented (corresponding to panel **J**, red frame).
All images shown in this figure have been contrast-optimized to facilitate the visualization of the examples, but the actual values were used when processing the images for quantification.

Purification and storage

Purification of iPSC-RPE cells was conducted using the same protocol described in a study previously reported [21]. When the RPE colonies reached an appropriate size, the cells were suspended in RPE maintenance medium and kept as a floating culture for about 10 days in a low cell adhesion plate (MS-90600Z, Sumitomo Bakelite Co., Ltd., Japan). Under the microscope, colonies consisting only of black RPE cells were selected. Then, they were transferred to 12-well plates coated with iMatrix, and cultured in RPE adhesion medium/RPE maintenance medium (1:1). Once the RPE cell colonies became attached to the dish, they were cultured in RPE maintenance medium with basic fibroblast growth factor (bFGF), which was changed every 2 to 3 days.

After 10–12 days of cell selection, unsuitable cells were removed, and the cells were passaged. The medium was aspirated and 1 mL of RPE washing solution was added and aspirated again for washing. Then, 0.5 mL of RPE washing solution was added and atypical cells were eliminated using micropipette tips under microscope observation. After the removal process, the medium was aspirated, 1 mL/well of PBS (-) was added and aspirated for washing, and then 0.5 mL of Trypsin-EDTA solution (203-20251, FUJIFILM Wako Pure Chemical Corporation, Japan) was added, followed by incubation at RT and 5% CO_2 for 8–10 min. Cells were detached by pipetting and collected into a 50 mL tube. After centrifugation (280 × *g*, 25 °C, 4 min), the supernatant was removed, and the pellet was resuspended in 1 mL/plate of RPE adhesion medium/RPE maintenance medium (1:1) and filtered through a 40 μm cell strainer (352340, Corning Incorporated, NY, U.S.A.). After counting the cells with a hemocytometer, the cells were seeded into 12-well plates. The medium was changed to RPE maintenance medium with bFGF.

After 1–3 days of cell passage, the medium was aspirated, the cells were washed with 0.5 mL of RPE maintenance medium, and 1 mL of RPE maintenance medium containing 10 ng/mL bFGF and 0.5 μM SB431542 was added. This medium was exchanged every 2–3 days.

The cells were stored when they formed hexagonal shapes after sufficient confluency. For that, the medium was aspirated, 1 mL/well of PBS (-) was added and then aspirated for washing, and 0.5 mL of Trypsin-EDTA solution (203-20251, FUJIFILM Wako Pure Chemical Corporation, Japan) was added, followed by incubation at 37 °C and 5% CO_2 for 10–15 min. After adding > 0.5 mL of RPE adhesion medium, the cells were detached using a cell scraper (MS-93100, Sumitomo Bakelite Co., Ltd., Japan). The cell suspension was filtered through a 40 μm cell strainer (352340, Corning Incorporated, NY, USA) and then centrifuged for 4 min at 280 × *g* to obtain a cell pellet. The pellet was resuspended in 1 mL of RPE adhesion medium/RPE maintenance medium (1:1) and filtered through a 40 μm cell strainer. After counting the cells with a hemocytometer, the cell suspension was centrifuged for 4 min at 280 × *g* to obtain a cell pellet. Then, STEM-CELLBANKER (CB047, Zenoaq

Resource Co., Ltd., Japan) was added until a cell concentration of 500,000 cells/0.5 mL/tube, and the cell suspensions were dispensed into cryovials. The cryotubes were placed in a cell freezing container at −80 °C for 3–24 h, and then stored at −150 °C.

Initiation of iPSC-RPE stock and recovery culture

Frozen vials of RPE cells were thawed in a 37 °C water bath and suspended in 4.5 mL of RPE adhesion medium. After centrifugation (280 × *g*, 25 °C, 4 min), the supernatant was removed and RPE adhesion medium/RPE maintenance medium (1:1) was added. After counting the cells with a hemocytometer, the cells were seeded into 24-well plates (0.5 mL/well).

After 1–3 days of cell seeding, the medium was aspirated, the cells were washed with 0.25 mL of RPE maintenance medium, and 0.5 mL/well of RPE maintenance medium containing 10 ng/mL bFGF and 0.5 μM SB431542 was added. This same type of medium was exchanged every 2–3 days.

Two weeks after seeding, the RPE cells were passaged. Two weeks after cell passage, the RPE cells were used for cell biological validation processes (RT-PCR, ELISA, and immunohistochemistry).

Validation — RT-PCR

Total RNA was extracted from transfected cells using RNeasy Micro Kit (#74004, QIAGEN, Germany). First-strand cDNA synthesis was performed on 500-1000 ng of total RNA, using SuperScript III (#18080-044, Thermo Fisher Scientific Inc., MA, USA) according to the manufacturer's instructions. Each mRNA transcript was amplified using PCR with the following primers:

BEST1 (+), 5′-dTAGAACCATCAGCGCCGTC
BEST1 (−), 5′-dTGAGTGTAGTGTGTATGTTGG
RPE65 (+), 5′-dTCCCCAATACAACTGCCACT
RPE65 (−), 5′-dCCTTGGCATTCAGAATCAGG
CRALBP (+), 5′-dGAGGGTGCAAGAGAAGGACA
CRALBP (−), 5′-dTGCAGAAGCCATTGATTTGA
GAPDH (+), 5′-dACCACAGTCCATGCCATCAC
GAPDH (−), 5′-dTCCACCACCCTGTTGCTGTA

Validation — ELISA

The collected media were centrifuged (90 × *g*, 4 °C, 1 min), and the supernatant was collected and stored at −80 °C. The amount of VEGF contained in the thawed medium was measured using the protocols and reagents from the VEGF Human ELISA Kit (BMS277-2, Thermo Fisher Scientific, USA), and the amounts of PEDF were measured using a Human ELISA Kit (RD191114200R, BioVendor, Czech Republic).

Validation — Immunohistochemistry

Cells were washed with PBS (-), fixed in 15% paraformaldehyde for 1 h at RT (approximately 25°C), and stored at 4°C after removal of PFA and addition of PBS (-). After removal of the solutions, cells were treated with 50 μL/well of 0.2% Triton X-100/PBS (-), incubated for 30 min at RT, washed with PBS (-), blocked with 50 μL of Blocking One (03953-95, Nacalai Tesque Inc., Japan), and incubated for 1 h at RT. After removal of the solutions, cells were stained at 4 °C o/n in 50 μL of the 1st antibody diluent (rabbit anti-ZO-1,

61-7300, Thermo Fisher Scientific Inc., MA, USA; anti-MITF, mouse anti-MiTF, ab80651, Abcam plc., Britain; antibody diluent, S2022, Agilent Technologies Inc., USA). After removal of the solutions, cells were washed with PBS (-) and then stained at RT for 1 h in 50 μL of the 2nd antibody diluent (Alexa Fluor 546 Goat Anti-mouse IgG, A-11030, Thermo Fisher Scientific Inc., MA, USA; Alexa Fluor 488 Goat Anti-rabbit IgG, A-11034, Thermo Fisher Scientific Inc., MA, USA; antibody diluent, S2022, Agilent Technologies Inc., USA) with DAPI (1 μg/mL, D1206, Thermo Fisher Scientific Inc., MA, USA). After removal of the solutions, cells were washed with PBS (-), and then 50 μL of PBS (-) was added. Images of immunohistochemistry samples were acquired using an IX73 inverted microscope (Olympus, Japan).

Bayesian optimization module

When no prior experimental results exist, the Bayesian optimization module generates the next query from random uniform sampling. When past experimental results are available, the Bayesian optimization module generates queries using two components: the Model updater and the Query generator (**Figure 2.10C**).

The Model updater updates the surrogate model to predict the experimental results given past experimental results: $D = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^{n}$. I adopted Gaussian process regression (GPR, **Figure 2.23**) with the ARD-RBF kernel as the surrogate model to estimate the expected score and confidence level for all unevaluated experimental parameters. Based on the experimental results shown in **Figure 2.7C**, the observation noise was assumed to follow a zero-mean Gaussian noise with a variance of 0.0039 at all points in the search space. By using the surrogate model, the Query generator generates the next queries in two steps. In step 1, the Query generator constructs an acquisition function that estimates the expected progress toward the optimal experimental parameter at a given experimental parameter x in the search space. I adopted the Expected improvement (EI) [30], a commonly used acquisition function in BO. EI estimates how much improvement over the current best score is expected from each point in the search space. In step 2, by using the acquisition function, the Query generator decides where to evaluate next, and our problem required the simultaneous performance of 48 experiments corresponding to 8 plates x 6 wells in each round. In addition, because the *DP* is a batch contextual parameter as described herein, a policy function that generates parameter sets taking such structural context into account must be incorporated. Therefore, I developed the Batch Contextual Local Penalization (BCLP) as a policy function to generate multiple points with context in parallel. The BCLP is a batch generation policy that extends the local penalization [31] to be applied to cases where complex structural context parameters exist. As shown in **Figure 2.10C**, for each value of the contextual parameter *DP* in ascending order, BCLP iteratively generated the parameter by maximizing and penalizing the acquisition function 48 times to obtain the next experimental parameters $\boldsymbol{X}_{\mathrm{next}}$ for each subsequent well (**Algorithms 1**, **2**). In addition, after each round, the more promising KP intervals were reconfigured by calculating the integral value of the acquisition function (**Algorithm 3**). I also replaced the queries that corresponded to the place of the top two pigmentation scores in the previous experiments with the parameter of the top two pigmentation scores in the previous experiments as a positive control.

Bayesian optimization for the iPSC-RPE protocol

In the batch Bayesian optimization for the iPSC-RPE differentiation protocol, I generated the next experimental queries for each round in four steps (**Algorithm 1**). First, I constructed a Gaussian process posterior using a past experimental dataset. When no prior experimental results existed, I generated the next query from random uniform sampling. Next, using the Gaussian process posterior, I obtained an acquisition function. In this study, I used the expected improvement [30] as the acquisition function. Third, using the acquisition function, I generated a next experimental parameter set $\boldsymbol{X}_{\mathrm{next}}$ using a policy function. In this study, I used batch contextual local penalization (BCLP) as the batch generation policy. Finally, after executing experiments on the parameter set $\boldsymbol{X}_{\mathrm{next}}$, I computed the optimal context for the Detach trypsin Period x_{DP} for the next round.

In Gaussian process regression, I assume the ARD-RBF kernel function k as follows:

$$k_{\mathrm{ARD-RBF}}\left(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\theta}\right) = \exp\left(-\frac{1}{2}\sum_{m=1}^{d}\left(\frac{\boldsymbol{x}-\boldsymbol{x}'}{l_m}\right)^2\right)$$

$\boldsymbol{\theta} = \{l_m\}_m^d$ represents d-dimensional length scale hyper-parameters that were optimized using the maximum-likelihood estimation in every GP fitting.

Based on the experimental results in **Fig. 2.7C**, the observation noise was assumed to follow a normal distribution with a variance of 0.039, at all points in the search space.

In this work , I used the expected improvement as an acquisition function. The expected improvement estimates how much improvement over the current best score y_{max} is expected from each one of the input parameters $\boldsymbol{x}$ in the search space, as shown in the following form:

$$EI(\boldsymbol{x}; D) = \left(y_{max} - \mu_n(\boldsymbol{x})\right)\Phi(Z) + \sigma_n(\boldsymbol{x})\phi(Z) \quad , \quad Z = \frac{y_{max}-\mu_n(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}$$

D is the set of the past experimental results. $\boldsymbol{\Phi}$ is the standard normal cumulative distribution function, and ϕ is the standard normal probability density function.

Input: The search space χ, GP prior (μ_0, σ_0, k), number of rounds M, number of Plates P, number of Wells W, Dataset $\mathcal{D}$= $\{\boldsymbol{x}_i, \mathrm{y}_i\}_{i=1}^n$

for t=1 to M do

1. Construct GP posterior (μ_t, σ_t, k) using $\mathcal{D}$.
2. Get the acquisition function $\alpha(\boldsymbol{x}; \mathcal{D})$.
3. Generate a experiment parameter set $\boldsymbol{X}_{next}$ using the policy function.

Execute the experiments $f(\boldsymbol{X}_{next})$

Append the experiment results to past data $\mathcal{D} = \mathcal{D} \cup \{(\boldsymbol{X}_{next}, f(\boldsymbol{X}_{next}))\}$.

4. Compute optimal context $\boldsymbol{c}_{DP}$ on Detatch trypsin Period in the next experiment.

end

Algorithm 1: Batch Bayesian optimization for the iPSC-RPE differentiation protocol

In the policy function (see **Figure 2.10C**), BCLP iteratively generated the parameter for each value of the contextual parameter *DP* in ascending order by maximizing and penalizing the acquisition function 48 times to obtain the next experimental parameters $\boldsymbol{X}_{\mathrm{next}}$. Here I show the algorithm that generates $\boldsymbol{X}_{\mathrm{next}}$ (**Figure 2.10C, Algorithm 2**). Starting from plate1, well1, BCLP fixes the value of the DP corresponding to a given well number and then maximizes the acquisition function to generate an experimental parameter for the well. The penalization of the acquisition function is then performed using a hammer function to the point where the experimental parameter is generated. Similarly, for the next

well, an experimental parameter is generated by maximizing the penalized acquisition function and is penalized using the hammer function. By repeating such maximization-penalization loops, BCLP generates the next experimental parameters for each subsequent well.

In **Algorithm 2**, the function $\varphi(\boldsymbol{x};\, \boldsymbol{x}_j)$ is a local penalizer of $\alpha(\boldsymbol{x};\, \mathcal{D}\,)$ at x_j such that:

$$\varphi\left(\boldsymbol{x}; \boldsymbol{x}_j, \hat{L}\right) = \tfrac{1}{2}\,\mathrm{erfc}(-z)$$

where

$$z = \frac{1}{\sqrt{2\sigma_n^2\left(\mathbf{x}_j\right)}}\left(\hat{L}\left\|\mathbf{x}_j - \mathbf{x}\right\| - \hat{M} + \mu_n\left(\mathbf{x}_j\right)\right)$$

for $erfc$ the complementary error function, $\hat{M} = \max_i\{\mathbf{y}_i\}$ is the best score in the past experiments and $\hat{L} = \max_{\mathcal{X}} \|\mu_\nabla(\boldsymbol{x})\|$ is an approximated Lipschitz constant. Both $\hat{M}$ and $\hat{L}$ were calculated in each round. I used $g(z) = z$ when the acquisition function $\alpha(\boldsymbol{x};\, \mathcal{D})$ was positive and the $g(z) = softplus(z)$ elsewhere [31].

Input: The acquisition function $\alpha(\boldsymbol{x};\, \mathcal{D})$, number of Plates P, number of Wells W
Output: The next experiment parameter set $\boldsymbol{X}_{next} = \{(\boldsymbol{x}_{t,p,w})\}_{(p,w)=1}^{(P,W)}$

1. Calculate utility functions from the acquisition function $\alpha(\boldsymbol{x},\, \mathcal{D})$
   $\tilde{\alpha}_0(\boldsymbol{x};\, \mathcal{D}) \leftarrow g(\alpha(\boldsymbol{x};\, \mathcal{D}))$
   $\tilde{\alpha}(\boldsymbol{x};\, \mathcal{D}) \leftarrow \tilde{\alpha}_0(\boldsymbol{x};\, \mathcal{D})$

2. Generate next experiment parameters $\boldsymbol{X}_{next} = \{(\boldsymbol{x}_{t,p,w})\}_{(p,w)=1}^{(P,W)}$ in Maximization-Penalization loop

for p=1 to P do
  for w=1 to W do
    1. Maximization-step: $\boldsymbol{x}_{t,p,w} \leftarrow \arg\max_{x\in\chi}\,\{\,\tilde{\alpha}\,(\boldsymbol{x};\, \mathcal{D})\,\}$
    2. Penalization-step: $\tilde{\alpha}(\boldsymbol{x};\, \mathcal{D}) \leftarrow \tilde{\alpha}_0(\boldsymbol{x};\, \mathcal{D}) \prod_{(k,h)=1}^{(p,w)} \varphi(\boldsymbol{x}\,;\, \boldsymbol{x}_{t,k,h},\, \hat{L})$
  end
end

Algorithm 2: The policy function for the iPSC-RPE differentiation protocol

For the contextual parameter x_{DP}, contexts on a fixed trypsin processing time $c_{DP,t,i}$ are assigned for each well depending on the round number t and the well number i, because of the implementation constraints of the protocol. After each round, the context on the x_{DP} is reconfigured as a variable. In the initial state (round $t = 1$), trypsin treatment time was assigned to the smallest well number (well 1) for eight minutes, and then it increased by three minutes for each additional well number.

$$\boldsymbol{c}_{DP,1} := (c_{DP,1,1}, c_{DP,1,2}, c_{DP,1,3}, c_{DP,1,4}, c_{DP,1,5}, c_{DP,1,6}) = (8, 11, 14, 17, 20, 23)$$

In the next round, the context on x_{DP} was allowed to move back and forth $\Delta c = 3$ in parallel while maintaining a three minute interval processing time between the different wells. Thus, the context in round 2 $\boldsymbol{c}_{DP,2}$ will be $\boldsymbol{c}_{DP}$ or $\boldsymbol{c}_{DP}^{-}$ or $\boldsymbol{c}_{DP}^{+}$ as follows:

$$\boldsymbol{c}_{DP} := \boldsymbol{c}_{DP,1} = (8, 11, 14, 17, 20, 23)$$
$$\boldsymbol{c}_{DP}^{-} := \boldsymbol{c}_{DP,1} - \Delta c = (5, 8, 11, 14, 17, 20)$$
$$\boldsymbol{c}_{DP}^{+} := \boldsymbol{c}_{DP,1} + \Delta c = (11, 14, 17, 20, 23, 26)$$

Using the flow shown in **Algorithm 3** in each round, I could choose whether the context in the next round $\boldsymbol{c}_{DP,t+1}$ will be $\boldsymbol{c}_{DP}$, $\boldsymbol{c}_{DP}^{-}$ or $\boldsymbol{c}_{DP}^{+}$, based on the past experimental data.

I performed a regression with GPR in each round and defined the acquisition function EI based on the regressions, deriving the integral value of the EI for x_{DP} for each different candidate interval of x_{DP} (either at 5 min, 8 min, or 11 min start) , respectively. When any derived integrals were improved by 5% or more compared to the integrals in the interval used in the current round, the interval for the *DP* for the next round was run in the interval that showed the largest integral value among the candidate intervals.

Input: The acquisition function $\alpha(\boldsymbol{x};\ \mathcal{D})$, current DP context $\boldsymbol{c}_{DP,t}$, context shift width Δc
Output: The next DP context $\boldsymbol{c}_{DP,\ t+1}$

1. Candidates of DP context ranges for the next round. (In this study, Δc= 3 min)

$\boldsymbol{c}_{DP} \leftarrow$ c$_{DP,t}$
$\boldsymbol{c}^{-}_{DP} \leftarrow$ c$_{DP,t}$ - Δc
$\boldsymbol{c}^{+}_{DP} \leftarrow$ c$_{DP,t}$ + Δc

2. Calculate values V, V^{-}, V^{+} that accumulate $\alpha(\boldsymbol{x},\ \mathcal{D})$ on each context ranges $\boldsymbol{c}_{DP}$, $\boldsymbol{c}^{-}_{DP}$, $\boldsymbol{c}^{+}_{DP}$

$V = \sum_i \int_{\chi} \alpha(\boldsymbol{x};\ \mathcal{D},\ x_{DP} = c_{DP,\ i})$
$V^{-} = \sum_i \int_{\chi} \alpha(\boldsymbol{x};\ \mathcal{D},\ x_{DP} = c^{-}_{DP,\ i})$
$V^{+} = \sum_i \int_{\chi} \alpha(\boldsymbol{x};\ \mathcal{D},\ x_{DP} = c^{+}_{DP,\ i})$

3. Calculate ratios R^{-}, R^{+} between each values defined above.

$R^{-} = V^{-}/V$
$R^{+} = V^{+}/V$

4. Choose the next DP context $\boldsymbol{c}_{DP,t+1}$ in following rules.

if (maxR^{-}, $R^{+} < 1.05$) then
  $\boldsymbol{c}_{DP,\ t+1} \leftarrow$ c$_{DP}$
end
else if $(R^{-} > R^{+})$ then
  $\boldsymbol{c}_{DP,\ t+1} \leftarrow$ c$^{-}_{DP}$
end
else if $(R^{-} \leq$ R$^{+})$ then
  $\boldsymbol{c}_{DP,\ t+1} \leftarrow$ c$^{+}_{DP}$
end

Algorithm 3: Detachment trypsin period adjustment on the iPSC-RPE differentiation protocol

---

Gaussian process (Bayesian Inference) Overview

In Bayesian optimization, the Gaussian process (GP) $p(f) = \mathcal{GP}(\mu, \sigma; k)$ is typically used as the surrogate model to approximate the I/O of the objective function f. The GP is characterized by three components: a mean function μ, a variance function σ, and a kernel function (positive-value covariance function) k. In case of past experimental data $D = \{(\boldsymbol{x}_i, y_i)\ \}_{i=1}^{n}$ of n observations, and Gaussian prior $\mathcal{GP}(0, \sigma; k)$, Gaussian posterior is then given in $\mathcal{GP}(\mu_n, \sigma_n; k)$, where

$$\mu_n(\boldsymbol{x}) = \boldsymbol{k}_n(\boldsymbol{x})^T \left(\boldsymbol{K}_n + \sigma^2 \boldsymbol{I}\right)^{-1} \boldsymbol{y}_n$$

$$\sigma_n^2(\boldsymbol{x}) = k(\boldsymbol{x}, \boldsymbol{x}) - \boldsymbol{k}_n(\boldsymbol{x})^T \left(\boldsymbol{K}_n + \sigma^2 \boldsymbol{I}_n\right)^{-1} \boldsymbol{k}_n(\boldsymbol{x})$$

$\mu_n(\boldsymbol{x})$ is the mean of Gaussian process posterior function, and $\sigma_n^2(\boldsymbol{x})$ is the variance of Gaussian process posterior function. In this work, I assume that the kernel function k is stationary. σ^2 is the observation noise of the experiment.

Figure 2.23 shows an example of a Gaussian process prior to being updated to a Gaussian posterior using observations. The uncertainty of GP distribution decreases around the observed points and increases further away from observations. For a more comprehensive explanation of Gaussian process regression, see [32].

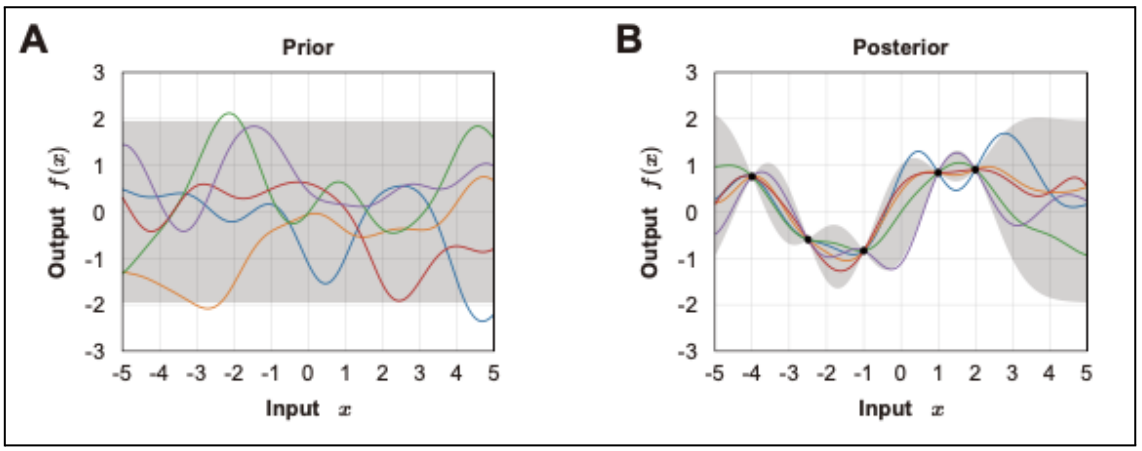


Figure 2.23: Demonstration of how Gaussian process regression updates a Bayesian posterior

Test of a Bayesian posterior updated in a Gaussian process (GP).

(**A**) Sample paths from the zero-mean GP prior.

(**B**) Sample paths from the GP posterior after executing some experiments (indicated by black dots).

The gray shaded area represents the pointwise mean plus and minus twice the standard deviation for each input value.

Bayesian optimization overview

Bayesian optimization is a common black-box optimization method used to determine the input parameter $\boldsymbol{x}$, which maximizes the objective function f with as few executions as possible using iterative experiment-observation loops.

$$\boldsymbol{x}_{\max} = \arg\max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})$$

Figure 2.24 shows a toy example of how sequential Bayesian optimization generates the next experimental parameter. Bayesian optimization iteratively generates the following experimental parameters in three steps: construction of a surrogate model (e.g., Gaussian process regression) for past experimental results (left column), definition of an acquisition function $\alpha(\boldsymbol{x}; D)$ (e.g., expected Improvement), and optimization of the acquisition function to obtain the next experimental parameter (right column). For a more comprehensive explanation and application of Bayesian optimization, see [33,34].

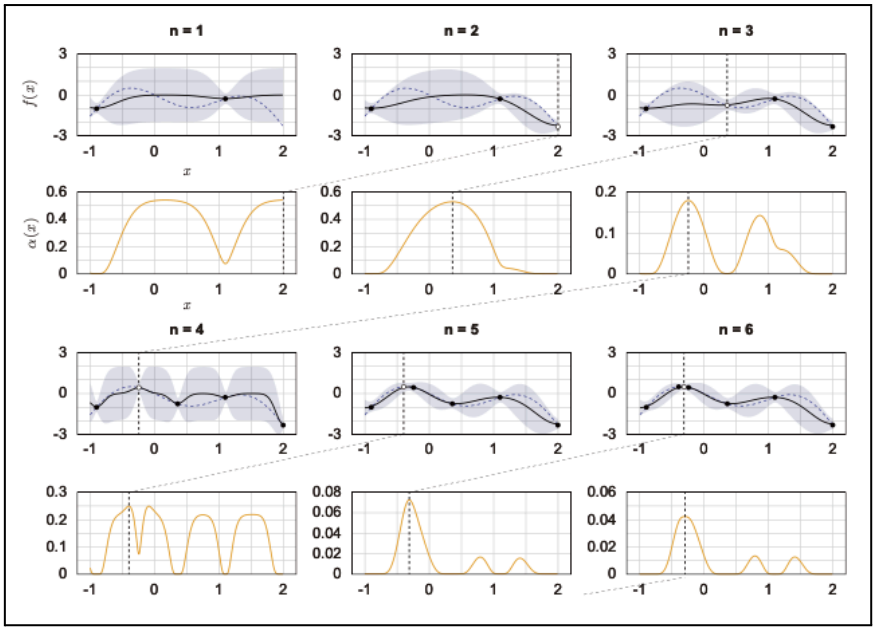


Figure 2.24: Demonstration of Bayesian optimization generation of experimental parameters

Sequential Bayesian optimization was tested by determining the maximum of a one-dimensional toy objective function. This figure illustrates the Bayesian optimization procedure over several iterations. The objective function is indicated by a purple dashed curve on the left side of each plot. The past experimental results are indicated by dots, and the newest experimental results are indicated by white dots. The GP posterior mean function is indicated by a black line, and covariance intervals are indicated by purple shaded areas. The plots on the right side show the acquisition functions in the orange curves. The value of the acquisition function is high where the model predicts a high objective (exploitation) and where the prediction uncertainty is high (exploration). The black vertical dashed lines show the place of acquisition max, a factor of importance to be tested in a subsequent experiment.

Statistical analysis

Statistical analyses were performed by Wolfram Mathematica version 11.2.0.0. In this study, $P < 0.05$ was considered significant (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and n.s. = not significant).

Data and code availability

All code that supports the findings of this study is available at https://github.com/labauto/LabDroid_optimizer. This code is based on GPyOpt [35].

Ethical approval

Human subjects: All experiments that involved the use of human-derived samples were reviewed and approved by the institutional review board of the Institutional Committee of RIKEN Kobe Branch (#Kobe1 2019-05 (3)).

References

1. Kitano, H. Artificial intelligence to win the Nobel Prize and beyond: Creating the engine for scientific discovery. *AI Mag.* **37**, 39–49 (2016).
2. Kitano, H. Nobel Turing Challenge: creating the engine for scientific discovery. *NPJ Syst. Biol. Appl.* **7**, 29 (2021).
3. King, R. D. *et al.* The automation of science. *Science* **324**, 85–89 (2009).
4. Nikolaev, P. *et al.* Autonomy in materials research: a case study in carbon nanotube growth. *Npj Comput. Mater.* **2**, 1–6 (2016).
5. Wigley, P. B. *et al.* Fast machine-learning online optimization of ultra-cold-atom experiments. *Sci. Rep.* **6**, 25890 (2016).
6. HamediRad, M. *et al.* Towards a fully automated algorithm driven platform for biosystems design. *Nat. Commun.* **10**, 5150 (2019).
7. Kino-Oka, M. *et al.* Automating the expansion process of human skeletal muscle myoblasts with suppression of myotube formation. *Tissue Eng. Part C Methods* **15**, 717–728 (2009).
8. Nishimura, A. *et al.* Fabrication of tissue-engineered cell sheets by automated cell culture equipment. *J. Tissue Eng. Regen. Med.* **13**, 2246–2255 (2019).
9. dos Santos, F. F., Andrade, P. Z., da Silva, C. L. & Cabral, J. M. S. Bioreactor design for clinical-grade expansion of stem cells. *Biotechnol. J.* **8**, 644–654 (2013).
10. Konagaya, S., Ando, T., Yamauchi, T., Suemori, H. & Iwata, H. Long-term maintenance of human induced pluripotent stem cells by automated cell culture system. *Sci. Rep.* **5**, 16647 (2015).
11. Thomas, R. J., Chandra, A., Hourd, P. C. & Williams, D. J. Cell Culture Automation and Quality Engineering: A Necessary Partnership to Develop Optimized Manufacturing Processes for Cell-Based Therapies. *JALA: Journal of the Association for Laboratory Automation* **13**, 152–158 (2008).
12. Liu, Y., Hourd, P., Chandra, A. & Williams, D. J. Human cell culture process capability: a

comparison of manual and automated production. *J. Tissue Eng. Regen. Med.* **4**, 45–54 (2010).

13. Soares, F. A. C. *et al.* Investigating the feasibility of scale up and automation of human induced pluripotent stem cells cultured in aggregates in feeder free conditions. *J. Biotechnol.* **173**, 53–58 (2014).
14. Matsumoto, E. *et al.* Fabricating retinal pigment epithelial cell sheets derived from human induced pluripotent stem cells in an automated closed culture system for regenerative medicine. *PLoS One* **14**, e0212369 (2019).
15. Ochiai, K. *et al.* A Variable Scheduling Maintenance Culture Platform for Mammalian Cells. *SLAS Technol* **26**, 209–217 (2021).
16. Kanie, K. *et al.* Effect of mechanical vibration stress in cell culture on human induced pluripotent stem cells. *Regen Ther* **12**, 27–35 (2019).
17. Kino-Oka, M. & Sakai, Y. Preface of the special issue 'Cell Manufacturability'. *Regen. Ther.* **12**, 1 (2019).
18. Kino-oka, M., Mizutani, M. & Medcalf, N. Cell manufacturability. *Cell Gene Ther. Insights* **5**, 1347–1359 (2019).
19. Regent, F. *et al.* Automation of human pluripotent stem cell differentiation toward retinal pigment epithelial cells for large-scale productions. *Sci. Rep.* **9**, 10646 (2019).
20. Yachie, N., Robotic Biology Consortium & Natsume, T. Robotic crowd biology with Maholo LabDroids. *Nat. Biotechnol.* **35**, 310–312 (2017).
21. Mandai, M. *et al.* Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. *N. Engl. J. Med.* **376**, 1038–1046 (2017).
22. Kuroda, T., Ando, S., Takeno, Y., Kishino, A. & Kimura, T. Robust induction of retinal pigment epithelium cells from human induced pluripotent stem cells by inhibiting FGF/MAPK signaling. *Stem Cell Res.* **39**, 101514 (2019).
23. Burger, B. *et al.* A mobile robotic chemist. *Nature* **583**, 237–241 (2020).
24. Gongora, A. E. *et al.* A Bayesian experimental autonomous researcher for mechanical design. *Sci Adv* **6**, eaaz1708 (2020).

25. Gulati, G. S. *et al.* Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
26. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2008).
27. Haruta, M. *et al.* In vitro and in vivo characterization of pigment epithelial cells differentiated from primate embryonic stem cells. *Invest. Ophthalmol. Vis. Sci.* **45**, 1020–1025 (2004).
28. Kawasaki, H. *et al.* Generation of dopaminergic neurons and pigmented epithelia from primate ES cells by stromal cell-derived inducing activity. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 1580–1585 (2002).
29. Osakada, F. *et al.* Toward the generation of rod and cone photoreceptors from mouse, monkey and human embryonic stem cells. *Nat. Biotechnol.* **26**, 215–224 (2008).
30. Jones, D. R., Schonlau, M. & Welch, W. J. Efficient Global Optimization of Expensive Black-Box Functions. *J. Glob. Optim.* **13**, 455–492 (1998).
31. González, J. I., Dai, Z., Hennig, P. & Lawrence, N. D. Batch Bayesian optimization via local penalization. *AISTATS* **51**, 648–657 (2015).
32. Williams, C. K. I. & Rasmussen, C. E. *Gaussian Processes for Machine Learning*. vol. 2 (MIT press Cambridge, MA, 2006).
33. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & de Freitas, N. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* **104**, 148–175 (2016).
34. Frazier, P. I. & Wang, J. Bayesian optimization for materials design. *arXiv [stat.ML]* (2015).
35. *GPyOpt: Gaussian Process Optimization Using GPy*. (Github).

Chapter3

3. Early and non-destructive prediction of the differentiation efficiency of human induced pluripotent stem cells using imaging and machine learning

3.1 Chapter Summary

The reproducibility and robustness of many directed differentiation protocols from human induced pluripotent stem cells (hiPSCs) remain low, and the long differentiation induction period significantly limits protocol optimization. To address this, I developed an early and non-destructive prediction system for the differentiation induction efficiency of hiPSCs using bioimage informatics. I employed a directed differentiation protocol for muscle stem cells (MuSCs), a promising cell source for the regenerative therapy of muscular dystrophy. Biological analyses suggested that days 14–38 are positive for forecasting the induction efficiency on day 82. Therefore, I conducted six independent experiments, inducing MuSC differentiation in a total of 34 wells, and captured a total of 5,712 phase contrast cell images between days 14 and 38. I selected fast Fourier transform (FFT) as the feature extraction method and confirmed that it captures the characteristics of cells during differentiation. By classifying images on each day using extracted features and machine learning, I found that samples with high and low induction efficiency could be predicted at approximately 50 days before the end of induction. This system is expected to contribute to regenerative therapy through effective protocol optimization.

3.2 Note on Contribution

Paper information: The content of this chapter has been published in the following journal article:

Hojo, M.A.*, **Tsuzuki, T.***, Ozawa, Y. et al. Early and non-destructive prediction of the differentiation efficiency of human induced pluripotent stem cells using imaging and machine learning. Sci Rep 15, 26821 (2025). (* The authors contributed equally to this work)

Clarification of Contributions: This study was conducted as an interdisciplinary collaborative project integrating stem cell biology and bioimage informatics. As a co-first author, I was primarily responsible for the conceptualization of the prediction system, experimental design regarding data acquisition, and the development of the machine learning framework. Specifically, my contributions are as follows:

1. Conceptualization

I proposed the core concept of the study: utilizing non-destructive phase-contrast imaging during the intermediate differentiation phase to predict the terminal yield of muscle stem cells (MuSCs), thereby addressing the issue of low reproducibility in long-term protocols.

2. Experimental Design

I designed the data acquisition scheme. Given the complexity of the 80-day culture protocol, the maximum feasible dataset was limited to 34 wells across six independent experiments. I defined the imaging strategy to maximize information within these experimental constraints.

3. Algorithm Development and Implementation

I determined that deep learning architectures were unsuitable for this study due to the limited sample size (34 biological replicates) inherent to the experimental design. To ensure robust generalization without overfitting, I strategically adopted a classical machine learning approach using fixed feature extraction.

4. Data Analysis and Visualization

I performed all computational analyses, including the evaluation of classifier performance using Area Under the Curve (AUC) and cross-validation to validate the model's reliability despite the small sample size. I also visualized the high-dimensional feature spaces using PCA plots to demonstrate how the extracted features correlated with differentiation efficiency and biological time points.

5. Manuscript Writing

I drafted the manuscript, specifically focusing on the sections describing the computational methodology, system architecture, and the interpretation of analytical results. I also contributed to the overall logic and discussion of the interdisciplinary findings.

3.3 Introduction

Human induced pluripotent stem cells (hiPSCs) are generated from adult somatic cells, and have infinite proliferation ability as well as the ability to differentiate into cells from any of the three germ layers [1,2]. Consequently, hiPSCs can be applied in a number of applications, such as regenerative therapy, drug discovery screening, disease modeling, and human developmental biology [2–6]. Currently, two strategies are used to support the differentiation of hiPSCs into specific somatic cells, namely direct reprogramming and directed differentiation [7]. In direct reprogramming, differentiated cells can be stably obtained in a short period via the forced expression of genes such as master transcription factors[7–9]. However, direct reprogramming is fundamentally unsuitable for clinical applications due to several limitations, including the risk of tumorigenesis, the generation of off-target cells, and the production of differentiated cells with extremely low maturation states[7,10,11]. Directed differentiation is a transgene-free method whereby cells are differentiated by culturing in a medium containing growth factors, differentiation factors, and small molecules, thus mimicking developmental processes [5,10,12–21]. This method carries a lower risk of tumorigenesis and often produces more mature differentiated cells compared to direct reprogramming methods[10]. Therefore, it is desirable to use cells and tissues differentiated via the directed differentiation method for regenerative therapy.

It is well known that the reproducibility and robustness of many directed differentiation protocols remain low. The induction efficiency varies in hiPSC clones, experimental batches, and even in the wells of plates due to the variability of the genetic and epigenetic state in each hiPSC clone[4,22–28], the researcher’s skill[29,30], slight changes in

seeding cell number, and bias of seeded cells in a culture dish. In addition, directed differentiation protocols for some cell types take several months for induction[5,10,16,17,20,21,31]. At present, there are very few methods that can be used to select samples with high induction efficiency at an early stage. This makes it difficult to elucidate important factors for improving and stabilizing induction efficiency. Moreover, assessing the properties of differentiated cells is laborious.

Therefore, novel methods are necessary to predict final differentiation induction efficiency at early time points. Although the assays to measure the expression of marker genes or proteins are currently common for evaluating cell conditions[10,13,32,33], they require skilled techniques, are time-consuming, costly, and most importantly, destructive to cells. Appropriate prediction methods must be non-destructive because the induction efficiency often varies in wells, even in the same experimental batch and condition. Phase contrast imaging is one of the most simple, low-cost, and non-destructive methods used to observe cell conditions. Recently, many studies have succeeded in estimating cell states from phase contrast cell images by combining imaging with machine learning[34–42]. Some examples include; detecting physical behavior of cells[34], determining whether cells are disease-related or not, determining the quality of differentiation, genetic interactions, gene and protein expression levels, and quantifying the number of days from the start of differentiation[35–42]. One limitation of these studies is that the properties of cells are judged at the time when the images are taken and these studies have not been able to predict the future cell condition. With regard to mesenchymal stem cells (MSCs) and hematopoietic stem cells (HSCs), some studies have succeeded in forecasting future differentiation states by utilizing computational image analysis or deep learning[43–45]. However, the number of cell lineages that can differentiate from MSCs and HSCs is significantly less than that from iPSCs. Furthermore, the differentiation induction period in these studies was up to 25 days, and prediction was performed between 18 hours and 7 days before. Some directed differentiation protocols from hiPSCs may even take several months for induction[5,10,16,17,20,21], which is currently a major bottleneck for the improvement of the reproducibility and robustness of these protocols. Therefore, to promote regenerative therapy, it is crucial to establish a method that is able to predict the differentiation state of hiPSCs at a very early stage using simple, low-cost, and non-destructive techniques, such as imaging.

In this study, I developed a system comprising phase contrast imaging and machine learning that is able to predict the final differentiation induction efficiency approximately 50 days prior to the induction period, which is less than half of the total induction period. I employed a directed differentiation protocol for the induction of muscle stem cells (MuSCs) from hiPSCs[10,46] that was developed by using muscle satellite cell marker MYF5-tdTomato reporter hiPSCs[10]. The previous report revealed that MYF5+ MuSCs have regenerative capacity for damaged muscle in Duchenne muscular dystrophy (DMD) model mice[10]; therefore, they are believed to be promising cell sources for regenerative therapy for muscular dystrophy. However, MuSC induction is a long-term process (approximately 80 days), with low reproducibility and robustness. By investigating gene and protein expression, I clarified that the expression of myogenic induction markers on day 38 were correlated with MuSC induction efficiency on day 82. From these results, I hypothesized that the final induction efficiency is predictable through cell imaging during myogenic induction phase from day 14 to 38. To test this hypothesis, and building upon observations of images acquired from day 14 to 38, I designed and validated a machine learning-based system to predict MuSC induction efficiency on day 82 (**Figure 3.1A**). Our system consists of two main processes: a Fast Fourier Transform (FFT)-based feature extraction process, followed by a classification process. In the feature extraction process, FFT is applied to each acquired phase contrast cell image to obtain its power spectrum. Subsequently, shell integration is performed on this power spectrum to generate a 100-dimensional, rotation-invariant feature vector, which is designed to capture morphological characteristics of the cells. In the subsequent classification process, these extracted feature vectors are utilized by a random

forest classifier to predict the MuSC induction efficiency on day 82 (**Figure 3.1B**). Our findings demonstrate that this system enables highly accurate prediction of samples with both high and low MuSC induction efficiency. Specifically, predictions made using images from days 31 or 34 for high-efficiency samples, and day 24 for low-efficiency samples, were particularly effective. In fact, classification using images from day 24 and day 34 resulted in a 43.7% reduction in the defective sample rate and a 72% increase in the number of good samples.

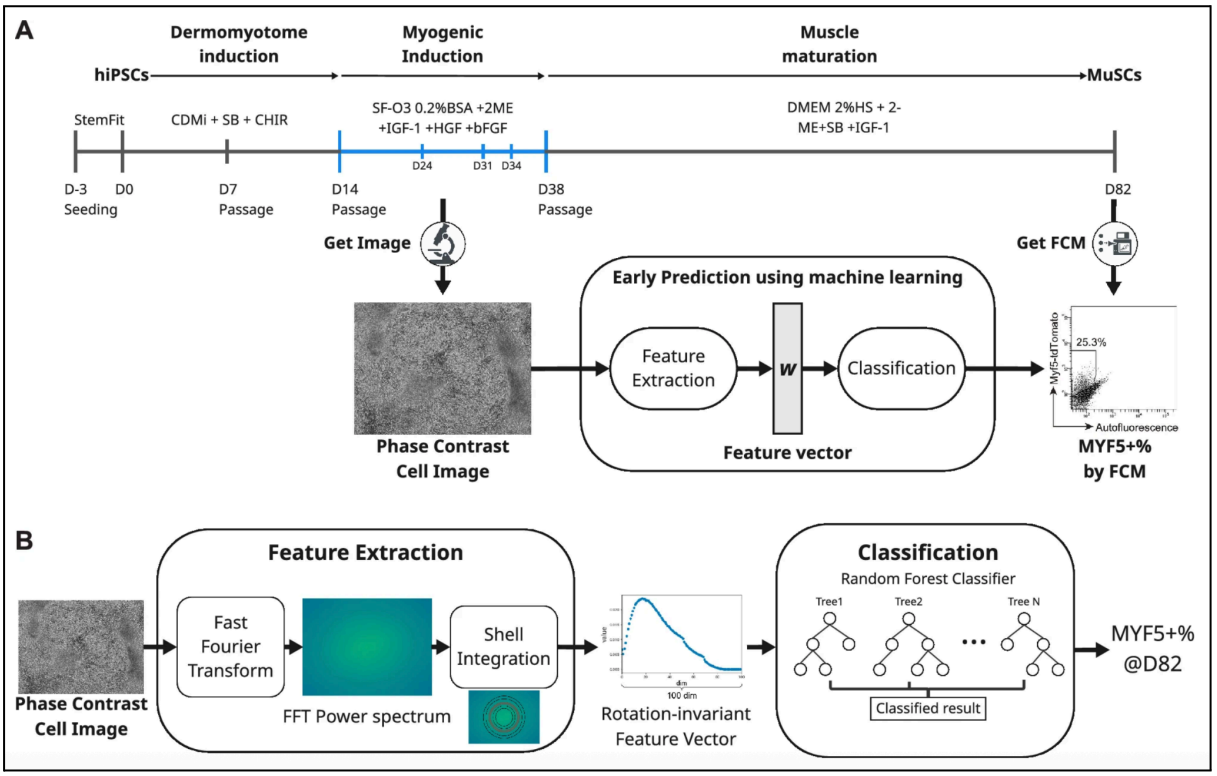


Figure 3.1: Schematic overview of the experimental workflow and the machine learning system for early prediction of muscle stem cell (MuSC) differentiation efficiency

(A) Schematic representation of the directed differentiation protocol for MuSCs and the general concept of the early prediction system using machine learning, which takes phase contrast images as input, performs feature extraction and classification to predict the final MYF5+% (day 82).
(B) Detailed workflow of the early prediction system. The Feature Extraction process begins with a phase contrast cell image. A Fast Fourier Transform (FFT) is applied to the image to obtain an FFT power spectrum. Subsequently, shell integration is performed on the power spectrum to generate a rotation-invariant feature vector, designed to capture morphological characteristics of the cells. This feature vector then serves as input for the Classification process. A Random Forest Classifier utilizes the extracted feature vector to predict the MYF5+% (day 82).

3.4 Results

3.4.1. The myogenic induction phase can be used to predict MuSC induction efficiency on day 82

In our MuSC induction protocol, dermomyotome cells were induced from hiPSCs via treatment with a Wnt agonist at a high concentration for 14 days. Next, the dermomyotome cells were treated with three types of growth factors; insulin-like growth factor 1 (IGF-1), hepatocyte growth factor (HGF), and basic fibroblast growth factor (bFGF) for 3 weeks in order to promote myogenic differentiation. The induced myotubes were then matured by switching the culture medium to a conventional muscle culture medium based on a low concentration of horse serum. MuSCs were obtained around day 80 (**Figure 3.1A**).

First, I investigated which phase of MuSC induction is suitable for prediction. *MYF5* has been reported to be expressed in satellite cells [47]. Previously, I generated MYF5-tdTomato reporter hiPSCs and induced differentiation into MuSCs using our protocol[10]. The resulting MYF5+ MuSCs expressed PAX7, a well-established marker of satellite cells[7], as well as other fetal muscle stem cell markers, and further demonstrated to have the ability to repair damaged muscle *in vivo*[10]. Therefore, previously generated MYF5-tdTomato reporter hiPSCs[10] were used to analyze the final MuSC induction efficiency via flow cytometry (FCM) on day 82. I performed quantitative real-time RT-PCR (qRT-PCR) for the representative genes on days 7, 14, and 38, and investigated the correlations to the MYF5 positive percentage (MYF5+%) on day 82. On day 7, the expression of an early mesoderm marker, *T* [48], a paraxial mesoderm marker, *TBX6* [49], and a dermomyotome marker, *SIX1* [50] were measured, and I observed no significant correlations to MYF5+% (**Figure 3.2A-C**). On day 14, the expression of dermomyotome markers, *DMRT2*, *PAX3*, and *SIX1* [32,50,51] were determined, and positive correlations were not observed (**Figure 3.2D-F**).

In contrast, significant positive correlations were detected between the expression of skeletal muscle markers, *MYH3* [52], *MYOD1* [53], and *MYOG* [54] on day 38 and MYF5+% on day 82 (**Figure 3.3A-C**). To confirm that these correlations were preserved at a protein level, immunocytochemistry (ICC) was performed. I found that samples with high myosin heavy chain (MHC) and MYOD1 expression tended to exhibit high MYF5+% on day 82 (**Figure 3.3D**). The MHC expressing area was quantified and was found to exhibit a significant positive correlation with MYF5+% on day 82 (**Figure 3.3E**). I also investigated whether MHC expression at day 38 correlated with the final MuSC differentiation induction efficiency in five additional cell lines. The differentiation induction efficiency in these cell lines was assessed using CDH13, a cell surface marker for MuSCs that I previously identified[53]. As a result, a significant positive correlation was observed between the MHC-positive area and the final CDH13 positivity rate (**Figure 3.2G**). These results suggest that monitoring the myogenic induction phase is promising for the prediction of MYF5+% on day 82.

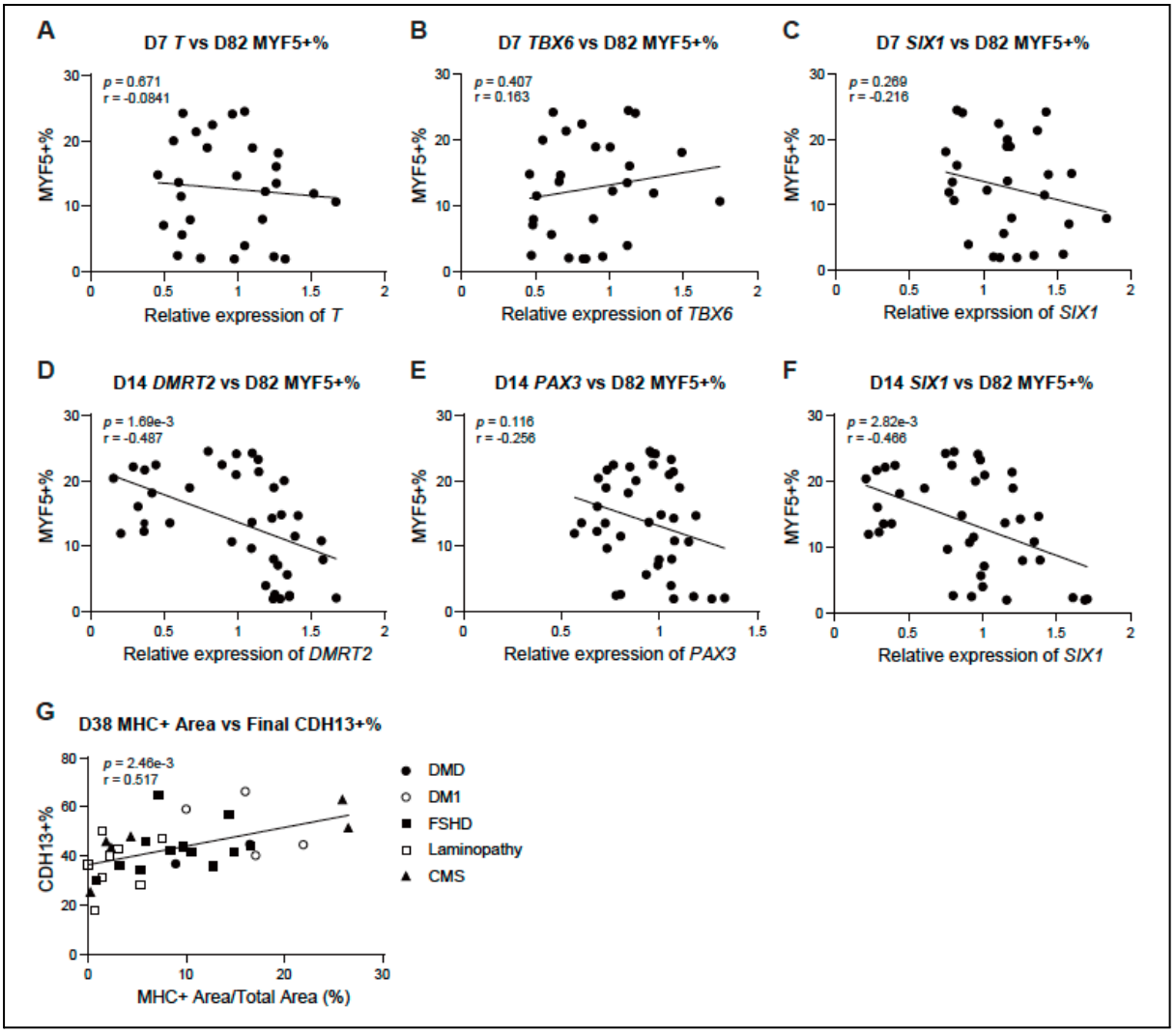


Figure 3.2: The dermomyotome induction phase is not suitable for the prediction of MuSC induction efficiency on day 82

(A-C) Correlation analysis between the MYF5+% score on day 82 and relative mRNA expression of T(A), TBX6 (B), or SIX1 (C) on day 7 (normalized to HPRT1). Data are pooled from 28 wells of experiments. The mRNA expression levels were examined in separate wells from those used for MYF5+% analysis and were cultured using the same experimental conditions.
(D-F) Correlation analysis between the MYF5+% score on day 82 and relative mRNA expression of DMRT2 (D), PAX3 (E), or SIX1 (F) on day 14 (normalized to HPRT1). Data are pooled from 39 wells of experiments. The mRNA expression levels were examined using separate wells from those used for MYF5+% analysis and were cultured using the same experimental conditions.
(G) Correlation analysis between the CDH13+% score on day 74–84 and MHC+ area calculated from images of ICC on day 38. Data are pooled from 32 wells across five patient-derived iPSC clones: Duchenne muscular dystrophy (DMD), myotonic dystrophy type 1 (DM1), facioscapulohumeral muscular dystrophy (FSHD) type 1, laminopathy, congenital myasthenic syndrome (CMS) with a GFPT1 mutation. Five images per well were taken at different positions.

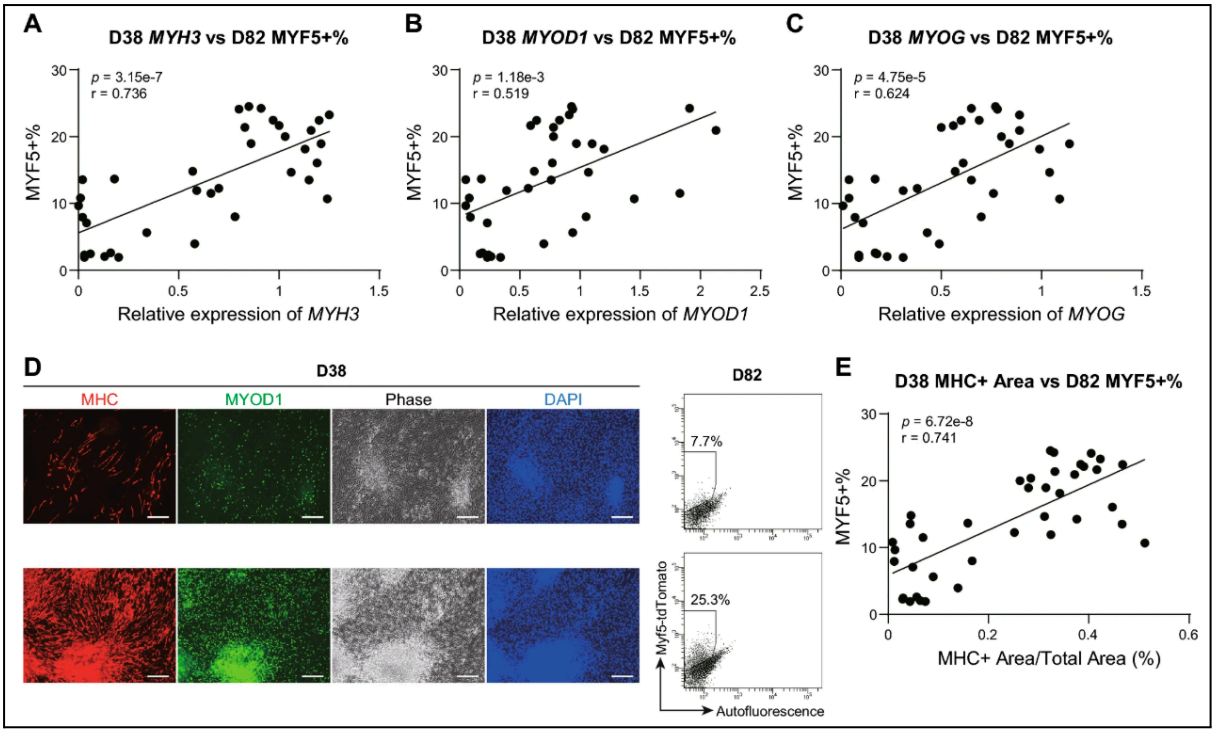


Figure 3.3: The myogenic induction phase can be used to predict muscle stem cell (MuSC) induction efficiency on day 82

(A-C) Correlation analysis between the MYF5+% score on day 82 and relative mRNA expression of MYH3 (B), MYOD1 (C), or MYOG (D) on day 38 (normalized to HPRT1). Data are pooled from 36 wells of experiments. The mRNA expression levels were examined using separate wells from those used for MYF5+% analysis and were cultured using the same experimental conditions.
(D) The results of immunocytochemistry (ICC) for myosin heavy chain (MHC) and MYOD1 on day 38, and flow cytometry (FCM) analysis for MYF5+% on day 82. Data are representative of 39 wells of experiments. ICC was performed using separate wells from those used for MYF5+% analysis and were cultured using the same experimental conditions. The scale bar represents 500 μm.
(E) Correlation analysis between the MYF5+% score on day 82 and MHC+ area calculated from images of ICC on day 38. Data are pooled from 39 wells of experiments. Between two to five images per well were taken at different positions.

3.4.2. The FFT-based feature extraction process was suitable for the prediction of MuSC induction efficiency from the myogenic induction phase

Based on the results displayed in **Figure 3.3**, phase contrast cell images obtained on eight separate occasions between day 14 and 38 were used for subsequent experiments. I performed six independent MuSC differentiation induction experiments, with 4–6 technical replicates per experiment, resulting in a total of 34 wells. For each well, 21 images were captured from different positions. In total, 5,712 images were acquired for MYF5+% prediction (**Figure 3.4A**). It was previously reported that the maximum MYF5+% score for this particular differentiation induction protocol was approximately 30 [54]. In our experiments,

this result was reproduced and the maximum MYF5+% score was 31. When purifying MuSCs with a cell sorter to evaluate MuSC properties, samples with a MYF5+% score over 20 can be efficiently collected, whereas samples with a MYF5+% score under 10 take a long time to sort and cells are often damaged. Therefore, I defined a high and low MYF5+% score as more than 20 and less than 10, respectively. Although deep learning has become common in image classification, it requires a substantial number of independent experiments for image collection, and our 34 experiments were insufficient [55]. Therefore, I decided to design a feature extraction process based on our observations of the images. By qualitatively comparing images with high and low MYF5+% scores, I found that cell density seemed to be higher in the samples with high MYF5+% scores between day 31 and 38 (**Figure 3.4B**).

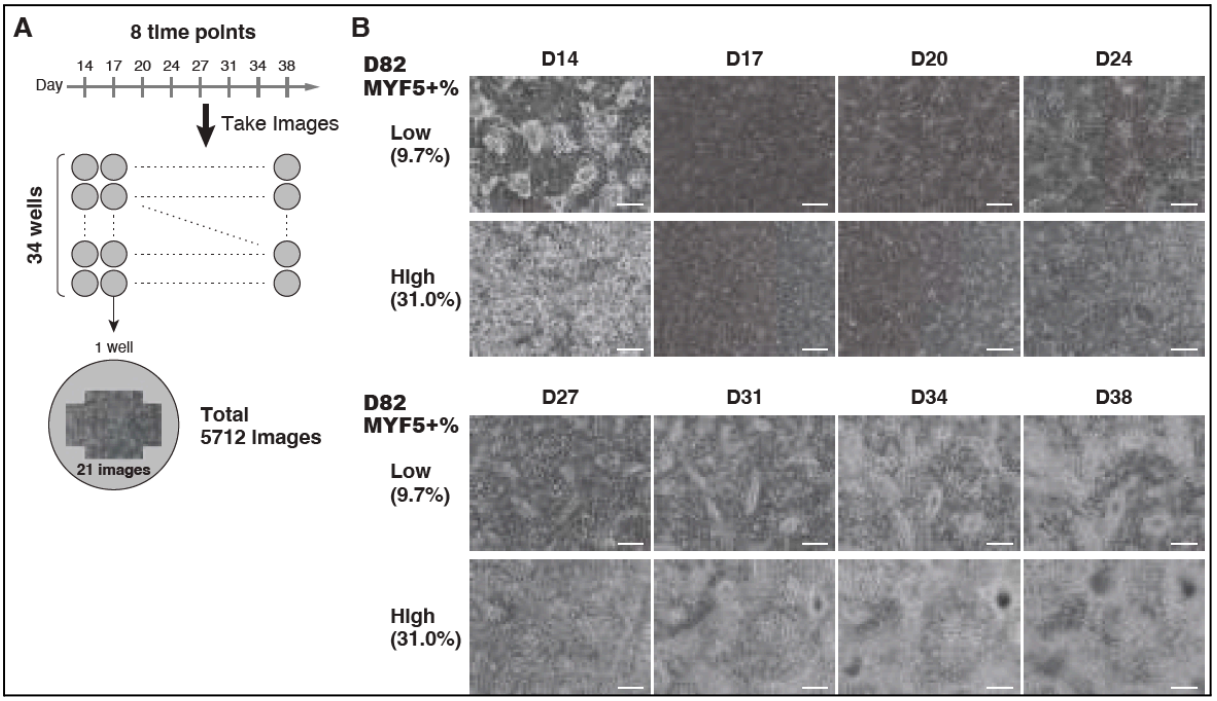


Figure 3.4: Schematic representation of the image acquisition procedure.

(A) Schematic representation of the image acquisition procedure. Images of each of the 34 wells were taken on eight separate occasions using 21 different positions. A total of 5,712 images were obtained.

(B) Phase contrast cell images from day 14 to 38. Images are representative of 21 different points of view in the well and 34 wells of experiments. The scale bar represents 500 μm.

Ruggeri et al. previously reported that a circular band in a FFT spectrum of an image contains frequency information related to cell density[56]; therefore, I designed a feature extraction process based on FFT. Our process comprised the following two steps: First, FFT was applied to a cell image to obtain a normalized FFT power spectrum. Then a shell integration was adopted to the normalized FFT power spectrum to obtain a rotation-invariant 100-dimensional feature vector (**Figure 3.1B**). To objectively confirm our observation that the MYF5+% score on day 82 correlated with the cell density between day 31 and 38, I compared the intensity of high-frequency components that should reflect a high cell density[56] using the images with high and low MYF5+% scores (**Figure 3.5A-C**). The intensity of high-frequency components was defined as a sum of the intensities of components above

the 30th feature vector. I investigated the images taken on days 31, 34, and 38 and found that the intensity was significantly higher in the group with a high MYF5+% score on days 34 and 38 (**Figure 3.5B-C**). On day 31, the difference between the high and low MYF5+% score groups was not significant, but the intensity of the high MYF5+% score group tended to be higher (**Figure 3.5A**). These results suggest that the cell density in the late myogenic induction phase was related to the MYF5+% score on day 82.

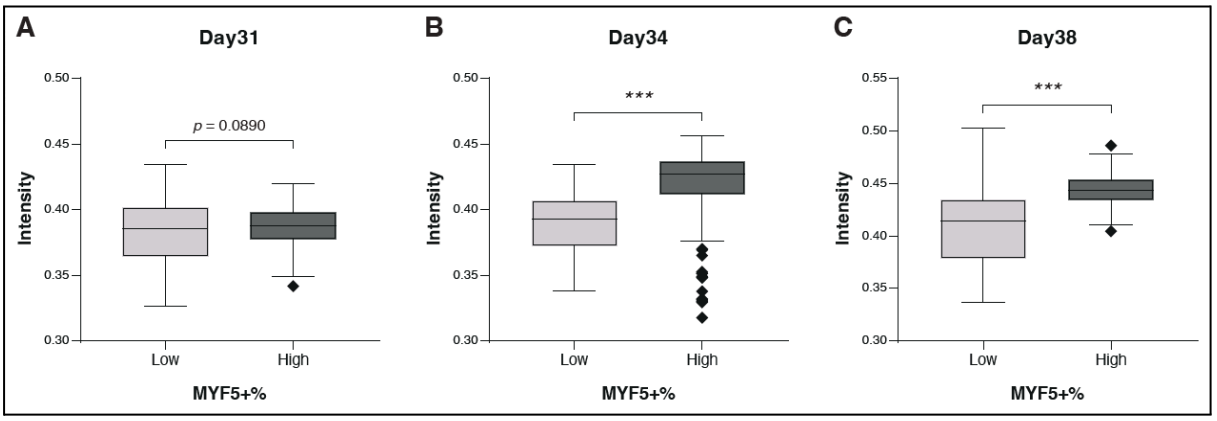


Figure 3.5: Comparison of the intensity of the high-frequency components between high and low MYF5+% score groups.

(A-C) Comparison of the intensity of the high-frequency components between high and low MYF5+% score groups on day 31 (A), day 34 (B), and day 38 (C). In each plot, the left bar represents the distribution of a low MYF5+% score group and the right bar represents the distribution of a high MYF5+% score group. Box-and-whisker plot is drawn in the style of Tukey. ***$p < 0.001$.

In addition, I obtained feature vectors for each of the 5,712 cell images and visualized them using principal component analysis (PCA) . Each cell image was grouped according to the date of differentiation, and the feature vectors for day 14 (only the dermomyotome induction phase) were plotted at a different point when compared to the other vectors of the myogenic induction phase (**Figure 3.6A**). This indicates that the FFT-based feature extraction process could capture the morphological changes that occur with differentiation. When comparing the results of the PCA analysis on day 34 and day 38 based on the MYF5+% scores (high: > 20%, middle: 10–20%, low: <10%), the dots in the high group appeared relatively clustered, whereas those in the middle and low groups were more dispersed (**Figure 3.6B-C**). These results suggest that the feature extraction method have the potential to capture morphological characteristics of cells associated with the MYF5+% score. Using the results displayed in **Figure 3.5-6**, I concluded that the use of the FFT-based feature extraction method was viable for the prediction of the MYF5+% score on day 82.

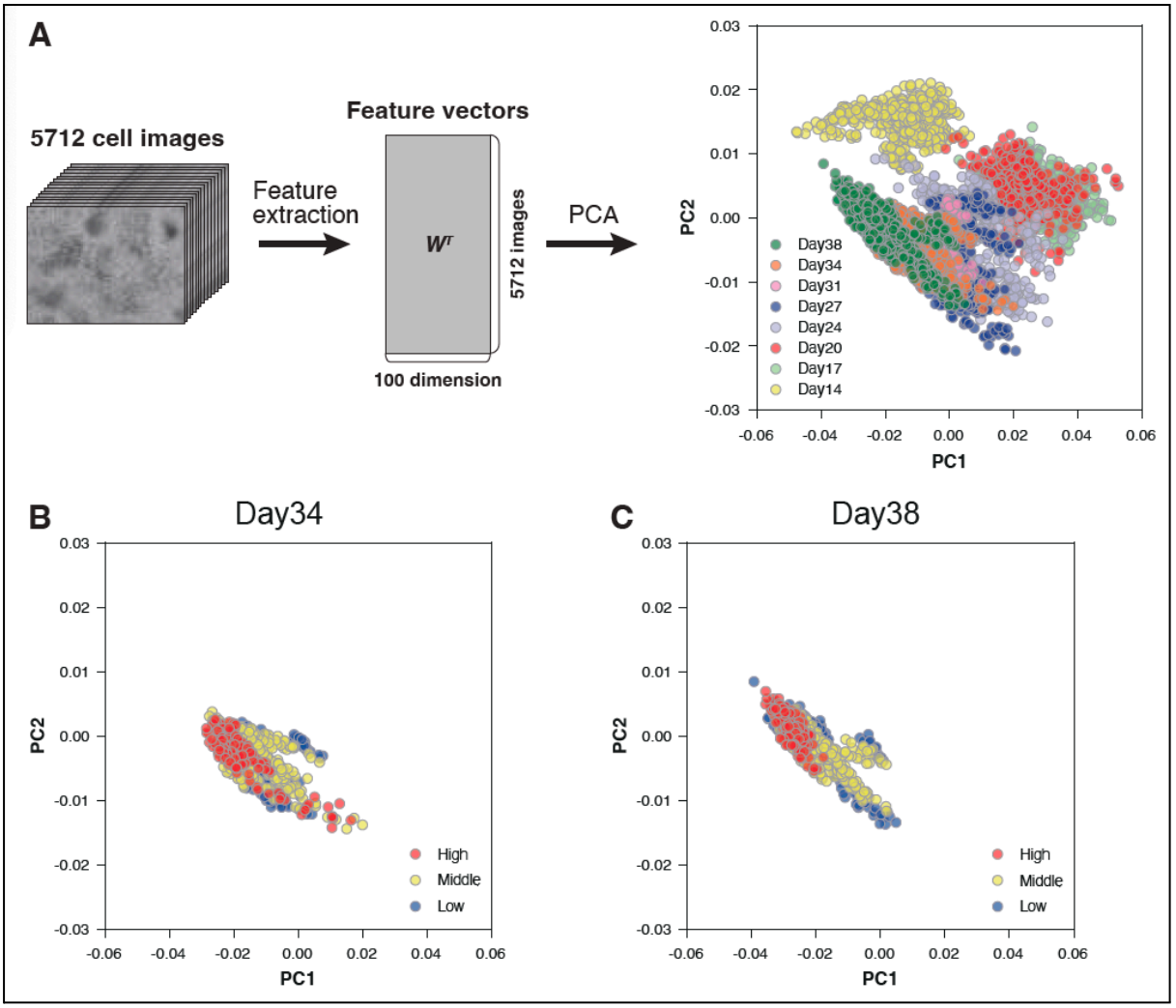


Figure 3.6: Comparison of the intensity of the high-frequency components between high and low MYF5+% score groups.

(A) Principal component analysis (PCA) for feature vectors from 5,712 images. The left side shows the feature extraction procedure, and the right scatter plot represents the PCA results. In the PCA plot, the horizontal axis represents the first component of PCA (PC1) and the vertical axis represents the second component of PCA (PC2). Each point in the scatter plot represents the value of each image. The colors of each point represent the day on which each image was taken.
(B-C) PCA results on day34 (B) and day38 (C). The colors of each point represent the MYF5+% score (high: > 20%, middle: 10–20%, low: <10%).

3.4.3. High and low MYF5+% scores were predictable approximately 50 and more days before the end of the induction period

In this study, I determined two classification tasks, Task 1 and 2, based on our definition of high and low MYF5+% scores. A 100-dimensional feature vector was extracted

using the FFT-based process. Cell images were classified using the extracted feature vectors and the random forest classifier[57], and I determined whether they met the requirements of Task 1 or 2 (**Figure 3.7A**). To determine which days were optimal for predicting high and low MYF5+% scores, I measured the classifier's performance using images from each day. For each date, the 34 wells of data were divided into either training data (25 wells) or test data (nine wells), and the classifier was trained using the training data (**Figure 3.7B**). The performance of the trained classifier was evaluated using test data and the area under the curve (AUC) was calculated. The AUC is the area under the receiver operating characteristic curve (ROC curve) and is generally used as an index of the performance of a classifier (**Figure 3.7C**)[58]. I performed four-fold cross-validation (CV) and obtained an average AUC (**Figure 3.7B**). Furthermore, I performed the same evaluation trial 50 times and changed the combination of train and test data division in each separate trial. Finally, distributions of the average AUCs for each day and Task were acquired (**Figure 3.7B**).

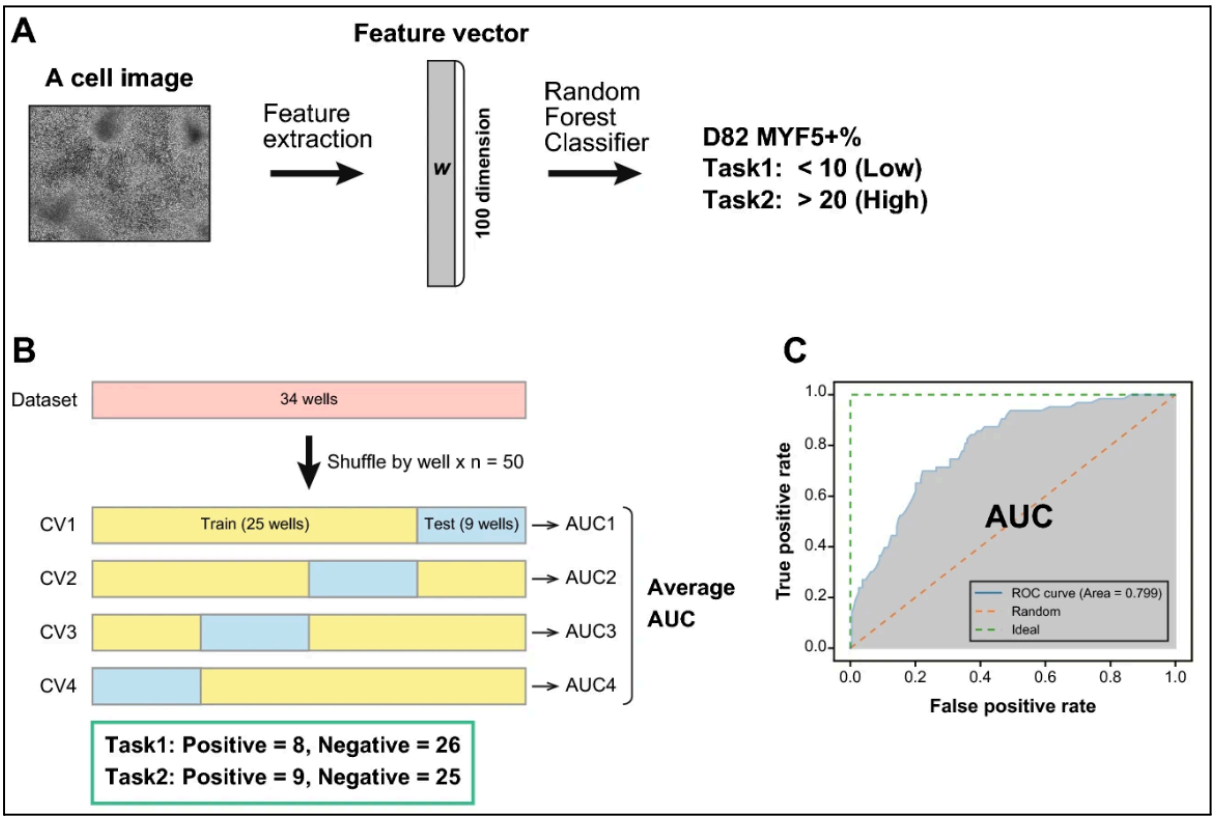


Figure 3.7: Schematic representation of the image classification procedure and testing structure.

(A) Schematic representation of the image classification procedure. Using the feature vectors extracted from the FFT, images were classified according to whether the MYF5+% score was less than 10 (Task 1) or more than 20 (Task 2) by a random forest classifier.
(B) Schematic explanation of the process used for the performance test of the classifier. Thirty-four wells of data were acquired and divided into either training data (25 wells) and test data (9 wells). Four-fold cross-validation (CV) was performed, and the average area under the curve (AUC) was calculated. The same trial was repeated 50 times; however, the combination of each data set was changed for each repeat. These results were used to calculate the distribution of average AUCs.

(C) Schematic explanation of the AUC. The AUC is the area under the receiver operating characteristic curve (ROC curve). The ROC curve is shown in blue. The AUC corresponds to the grey-filled region.

In Task 1 (Low: MYF5+% score < 10), the average AUCs on day 24 were significantly higher than those on the other days (median: 0.76), and they decreased as the induction period progressed (**Figure 3.8A**). In Task 2 (High: MYF5+% score > 20), the average AUCs were significantly higher on days 31 and 34 (median: 0.86), followed by day 24 (**Figure 3.8B**). I concluded that these specific days are appropriate for the prediction of samples with high and low induction efficiency. The results displayed in **Figure 3.8A-B** indicate that features correlating to high and low MuSC induction efficiency were included in the images obtained on days 31 or 34, and 24 respectively. To consider features other than cell density that may affect the performance of the classifier, I monitored the expression of MHC, which is the structural protein of myotubes, using ICC during the myogenic induction phase (days 17–38). I found that MHC expression began around day 24 and was strongly expressed from day 31 (**Figure 3.9**). Based on these results, I consider that morphological changes associated with MHC expression may have been one of the features utilized by the classifier in addition to cell density.

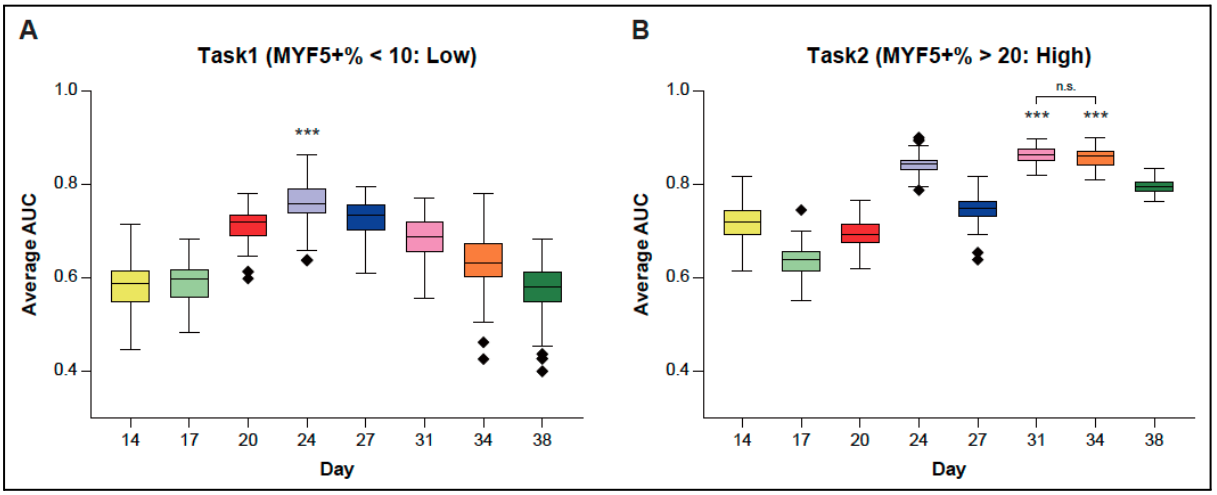


Figure 3.8: High and low MYF5+% scores were predictable approximately 50 or more days before the end of the induction period

(A-B) Performance tests of the classifier for Task 1 (A) or Task 2 (B). The horizontal axis represents the number of days from the start of differentiation. The vertical axis represents the average AUC value. Data are pooled from each average AUC derived from the 50 performance tests. The box-and-whisker plot is drawn in the style of Tukey. ***q < 0.001 to all other days. n.s. not significant.

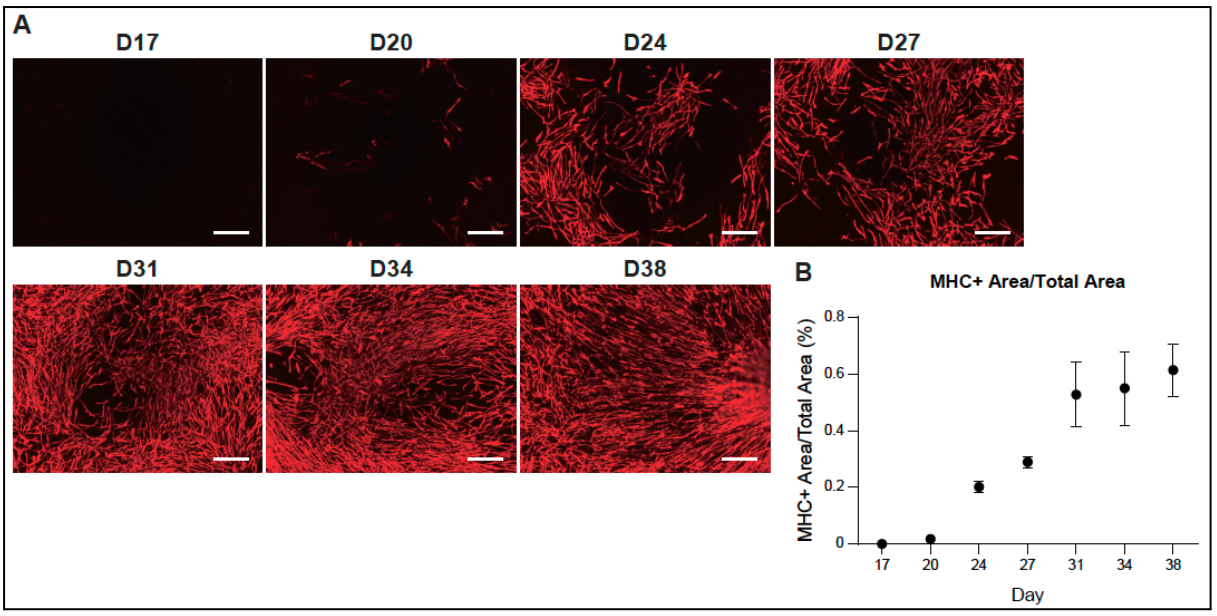


Figure 3.9: The possibility that the classifier also used morphological change related to MHC expression for prediction.

(A) The results of ICC for MHC from day 17 to 38. The scale bar represents 500 μm.
(B) MHC+ area calculated using images of ICC between day 17 and 38. Four images per well were obtained at different positions of the well.

Next, I evaluated the performance of the classifier at day 24 for Task 1 and day 34 for Task 2 after applying the classifier as a filter to improve yield in the manufacturing process. The evaluation results are shown in **Figure 3.10**. For Task 1, which detects and excludes images with MYF5+% scores below 10% (defective products) as positive, I assessed the performance of the classifier using two metrics: False Positive Rate (FPR) and Impurity Change. The FPR indicates the proportion of images that actually have MYF5+% scores above 10% but were incorrectly detected as below 10%. Here, the FPR was 0.19, indicating that 20% of the images, which were not defective, were erroneously classified as defective. Impurity Change measures the change in the proportion of images with MYF5+% scores below 10% in the dataset before and after classification. Initially, among the 714 cell images, 147 (20.6%) had MYF5+% scores below 10%. After classification, 196 images were excluded as defective, leaving 518 images. Of these, 60 images (11.6%) had MYF5+% scores below 10%, resulting in an Impurity Change of 0.116/0.206 = 0.563. This significantly reduced the defect rate by 43.7% ($p = 2.91 \times 10^{-20}$) (**Figure 3.10A**). For Task 2, which detects and retains images with MYF5+% scores above 20% (high-quality products) as positive, I evaluated the classifier's performance using False Negative Rate (FNR) and Impurity Change. The FNR indicates the proportion of images that actually have MYF5+% scores above 20% but were incorrectly detected as below 20%. Here, the FNR was 0.078, indicating that 7.8% of the high-quality images were mistakenly classified as lower quality. Impurity Change measures the change in the proportion of images with MYF5+% scores above 20% in the dataset before and after classification. Initially, among the 714 cell images, 189 (26.5%) had MYF5+% scores above 20%. After classification, 383 images were retained as high-quality. Of these, 174 images (45.5%) had MYF5+% scores above 20%, resulting in an Impurity Change of 0.455/0.265 = 1.719. This significantly concentrated the high-quality product rate by 72% ($p = 1.10 \times 10^{-39}$) (**Figure 3.10B**).

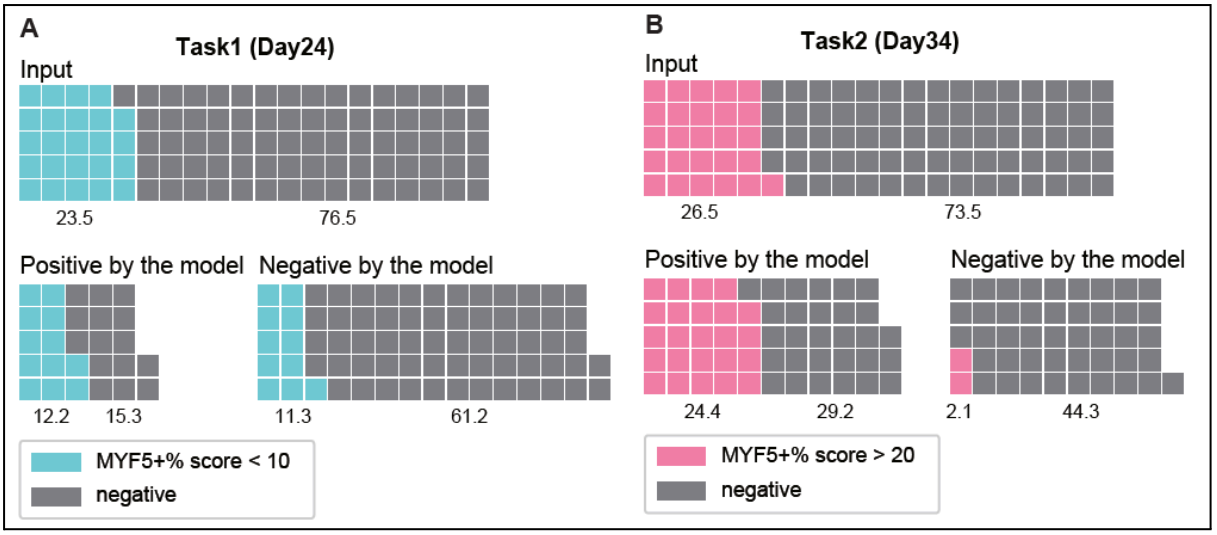


Figure 3.10: The possibility that the classifier also used morphological change related to MHC expression for prediction.

(A-B) Waffle plots illustrating the classifier's performance when used as a filter in Task 1 (A) and Task 2 (B). Each plot contains 100 squares, with each square representing 1% of the total cell images, normalized for visual clarity. Colored boxes indicate positive samples for each task (Task 1: MYF5+% < 10%, Task 2: MYF5+% ≥ 20%). “Input” shows the initial distribution of samples. “Positive by the model” shows the distribution of samples classified as positive by the classifier. “Negative by the model” shows the distribution of samples classified as negative by the classifier.

Based on these results, I concluded that the classifiers I constructed for Task 1 and Task 2 were useful for application scenarios. In particular, the classifier for Task 2 demonstrated high effectiveness.

3.5 Discussion

Directed differentiation from hiPSCs is suitable for regenerative therapy because it has a low risk of tumorigenicity. Currently, there are no available methods that can be used to predict final differentiation induction efficiency at an early stage, and at present, many protocols take several months for induction, with low reproducibility and robustness. In this study, I succeeded in predicting the future differentiation induction efficiency using phase contrast cell images in the early phase of differentiation by designing an image classification process with FFT and a random forest classifier based on biological investigation and visual image comparison. I also demonstrated the performance of our classifier, showing that it is particularly effective in enriching samples that exhibit high differentiation efficiency.

Our directed differentiation protocol for MuSCs consisted of three phases and took approximately 80 days to complete (**Figure 3.1**). First, to investigate which phase could be used to predict the final induction efficiency, I analyzed the expression of representative genes in each phase and determined their correlation to the final differentiation induction efficiency (**Figure 3.2-3**). I found significant positive correlations between the expression of myogenic-related genes and proteins on day 38 and the MYF5+% score on day 82 (**Figure**

3.3). These results suggest the following: First, in our protocol, MuSCs were induced with other muscle lineage cells [57]. Therefore, the samples with a high myogenic induction efficiency on day 38 were considered to have high differentiation efficiency of whole myogenic-lineage cells, including MuSCs. Second, the MuSCs induced using our protocol exhibited a fetal muscle stem cell-like phenotype [10], and a previous study has reported the importance of myofibers as a niche of muscle stem cells *in vivo* [59]. Therefore, it is presumed that the high degree of differentiation of myotubes on day 38 promoted the differentiation of the MuSCs. To improve the robustness of the differentiation induction of MuSC, it could be crucial to optimize the protocol of the myogenic induction phase.

Based on our qualitative observation, I developed a process that extracts image features using a FFT and predicts high and low MYF5+% scores by processing the features via a random forest classifier (**Figure 3.1B, Figure 3.7A**). Surprisingly, it was revealed that I could predict what the high and low MYF5+% scores would be on day 82 with high accuracy on days 31 or 34, and 24, respectively, which is almost 50 days prior to the end of the differentiation induction period (**Figure 3.8**). Furthermore, it was shown that images on day 24 and day 34 contributed to the elimination of bad samples and the selection of good samples (**Figure 3.9**). Until now, many studies have reported estimate cell states, such as the gene expression level, using cell imaging combined with machine learning [35–42]. However, in these reports, the cell states were judged only at the time of the image being taken. Some studies on the differentiation of MSCs and HSCs have forecasted future differentiation states using cell images. Kazemimoghadam et al. identified a suitable prediction point that can be used to judge high neural differentiation potential using computational image analysis [5,10,16,17,20,21]. Buggenthin et al. and Matsuoka et al. successfully predicted future hematopoietic lineage choice and osteogenic differentiation potential using cell imaging combined with deep learning [54]. However, the induction periods used in these studies were all shorter than 1 month, which is substantially shorter than many directed differentiation protocols for hiPSCs [5,10,16,17,20,21]. **Furthermore, prediction points were only obtained between 18 hours and 7 days before the end of the induction period. In contrast, in the present study, I succeeded in predicting the differentiation induction efficiency after 82 days of culture using images taken approximately 50 days prior, which was during the early phase of differentiation induction. This could contribute significantly to the improvement of reproducibility and robustness in differentiation induction.** In addition, iPSCs have greater pluripotency than MSCs and HSCs, and this study demonstrated the feasibility of cell image-based future prediction for differentiated cells derived from hiPSCs. It is important to develop early-stage prediction methods for other hiPSC-derived differentiation induction systems in order to widely promote regenerative therapy.

A previous study conducted by Ruggeri et al. demonstrated that FFT can be used to obtain information on cell density [56]. **Figure 3.5** demonstrated that the high-frequency component was significantly higher in the high MYF5+% score group on days 38 and 34, whereas no significant difference was observed on day 31. However, the accuracy for the prediction of both high and low MYF5+% scores was substantially higher on day 31 than on day 38 (**Figure 3.8**). Furthermore, it was clarified that images on day 24 were more suitable for the prediction of low MYF5+% scores (**Figure 3.8B**) even though no big difference in cell density was observed on day 24 (**Figure 3.8A, Figure 3.11**). These results suggest that, in addition to cell density-related features, the random forest classifier used some morphological characteristics for image classification. In **Figure 3.9**, I indicated that the day on which MHC expression was largely increased corresponded with the day with high prediction accuracy. Therefore, it is suggested that morphological changes related to MHC expression were one of the features used for prediction.

By adopting our proposed method, I succeeded in constructing a classifier that can accurately predict the final MYF5+% scores (approximately 0.8 AUC) and concentrate

samples with high differentiation efficiency (**Figure 3.8**). However, some samples were not classified correctly by our classifier. One possible approach that can be used to improve the classifier's performance is to increase the number of experiments conducted to obtain samples for training. As more data is accumulated from experiments, I expect to improve the classifier's performance by extracting more informative features from images using more flexible methods such as deep learning.

In this study, I demonstrated the potential to eliminate samples with failed differentiation and enrich those with successful differentiation around days 24–34. Although the culture period has been significantly reduced compared to the conventional ~80 days, achieving an even earlier prediction would be more desirable from the perspective of minimizing cell culture labor and expenses. In image analysis, I used the images on day 14 as the control because day 14 fell within the dermomyotome induction phase , and the expression of representative genes on day 14 were not correlated with the MYF5+% score on day 82 (**Figure 3.2**). Nevertheless, the average AUCs for the prediction of a high MYF5+% score on day 14 were relatively high (approximately 0.7 AUC) (**Figure 3.8B**). This suggests that some important factors for MuSC induction could be detected even on day 14 if I developed a higher-resolution image analysis method or another non-destructive method.In conclusion, I developed a system that is able to predict and select the samples with high and low induction efficiency of MuSCs using phase contrast cell images. Conventionally, the determination of the cell condition in the differentiation process has required an expert's eye. Using this system, I can determine the final induction efficiency with consistency approximately 50 days before the induction period is complete, which is less than half of the total induction period. This could contribute to elucidating important factors that can be used to improve the reproducibility and robustness of differentiation induction; thus, reducing the cost and time for experiments. Furthermore, this is expected to promote the evaluation and leverage of differentiated cells through their effective production, and I believe that this system may accelerate hiPSC-based regenerative therapy. As a next step, I should focus on testing whether a similar method could be applicable to another directed differentiation protocol.

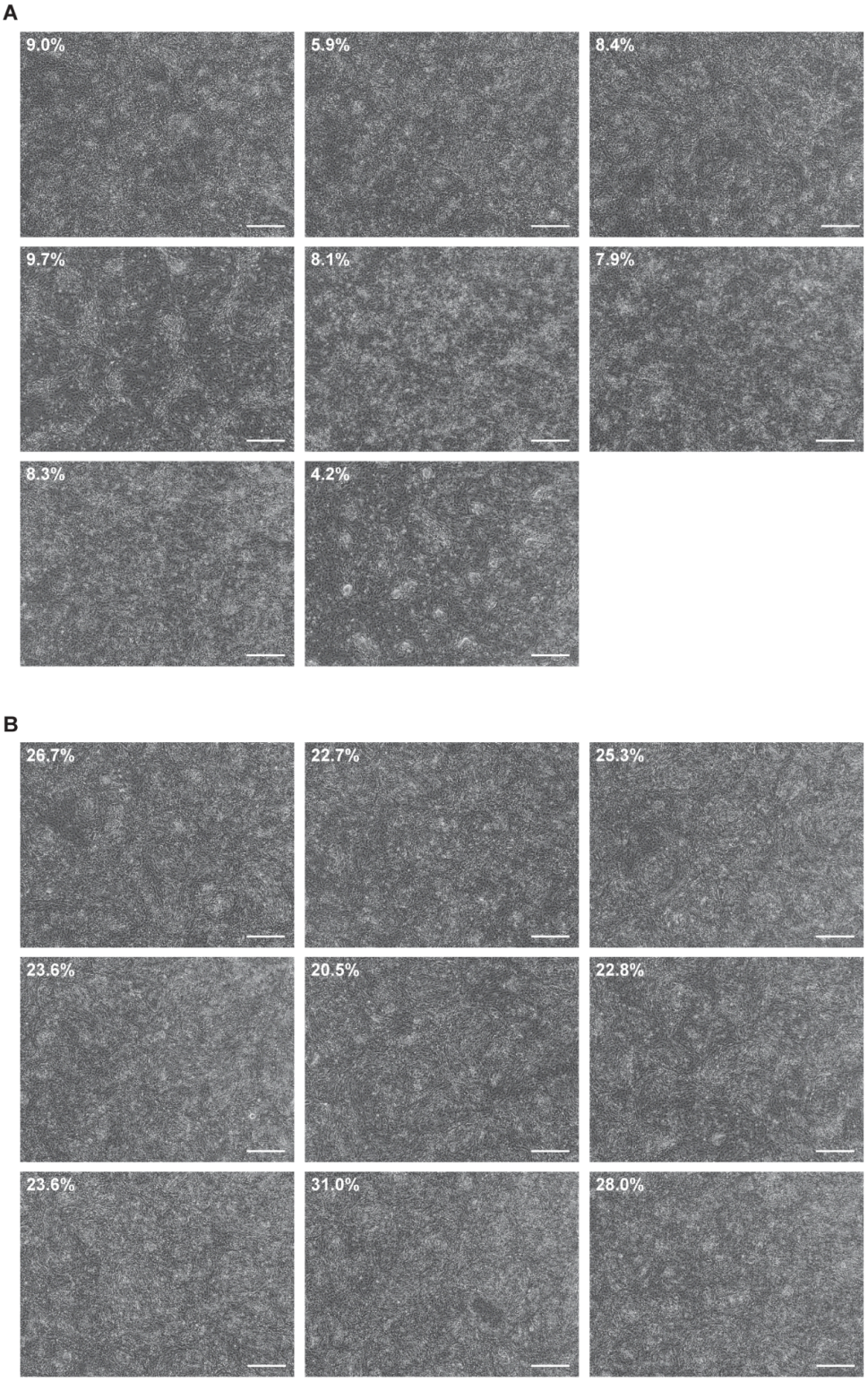


Figure 3.11: Phase contrast cell images on day 24

(A) Images from all samples that had a MYF5+% score of less than 10 on day 82. Images are representative of 21 points of view per well. The scale bar represents 500 µm.
(B) Images from all samples that had a MYF5+% score of more than 20 on day 82. Images are representative of 21 points of view per well. The scale bar represents 500 µm.

3.6 Methods

hiPSC line and maintenance culture

The hiPSC clone 201B7-derived[10] MYF5-tdTomato reporter line was used for MuSC differentiation induction. The original 201B7 cell line distributed by RIKEN BRC has been confirmed to have a normal karyotype through karyotypic analysis (https://cellbank.brc.riken.jp/cell_bank/CellInfo/?cellNo=HPS0063&lang=En). The MYF5-tdTomato reporter line was previously established by a CRISPR-Cas9-based knock-in system [59]. hiPSCs were cultured and maintained under feeder-free culture conditions as previously described [60]. Briefly, cells were passaged at 2,000–10,000 cells/well in 6-well plates coated with iMatrix-511 (Nippi, Incorporated, Japan) once a week and maintained in 2 mL/well of StemFit AK02 medium (Ajinomoto, Japan). The medium was replaced three times per week with fresh StemFit AK02 medium. I used five patient-derived iPSC clones: A clone derived from a Duchenne muscular dystrophy (DMD) patient (Clone ID: CiRA00111, which was referred to as DMDΔ44-1 in our previous report[60]), a clone derived from a myotonic dystrophy type 1 (DM1) patient (Clone ID: CiRA00211, which was referred to as Patient-2 in our previous report[61]), a clone derived from a facioscapulohumeral muscular dystrophy (FSHD) type 1 patient (Clone ID: CiRA00655, which was referred to as F1#1 in our previous report[62]), a clone derived from a laminopathy patient (Clone ID: CiRA00457, which was referred to as LMNA in our previous report[63]), and a clone derived from a congenital myasthenic syndrome (CMS) patient who had a mutation in *GFPT1* gene (c.722–723insG) (Clone ID: M16#66). All patient-derived iPSC clones were established by overexpression of Oct3/4, Sox2, L-Myc, Klf4, and Lin28 by using episomal vectors.

Directed differentiation of MuSCs from hiPSCs

Differentiation induction of MuSCs was performed following our previously reported protocol[10,46]. hiPSCs were dissociated with Accutase (Nacalai tesque, Japan) and plated at 5,000–20,000 cells/well on Matrigel (Corning, NY, USA)-coated 6-well plates in 2 mL/well of StemFit AK02 medium supplemented with 10 μM Y-27632 (Nacalai tesque). After 2 days, the medium was replaced with StemFit AK02. On the following day, the medium was switched to a chemically defined medium (CDMi) supplemented with 5 μM SB431542 (SB, Nacalai tesque or FUJIFILM Wako Pure Chemical Corporation, Japan) and 10 μM CHIR99021 (CHIR, FUJIFILM Wako Pure Chemical Corporation), and dermomyotome induction was initiated (day 0). CDMi is composed of Iscove's Modified Dulbecco's Medium (IMDM, FUJIFILM Wako Pure Chemical Corporation, 098-06465) and Ham's F-12 (FUJIFILM Wako Pure Chemical Corporation, 087-08335) at a 1:1 ratio and is supplemented with 1% bovine serum albumin (BSA, Sigma-Aldrich, STL, USA, A9418), 1% Penicillin-Streptomycin Mixed Solution (Nacalai tesque), 1% CD Lipid Concentrate (Thermo Fisher Scientific, Inc., MA, USA), 1% Insulin-Transferrin-Selenium (Thermo Fisher Scientific, Inc.), and 450 μM 1-Thioglycerol (Sigma-Aldrich). On days 3 and 5, the medium was replaced with the same medium that was used on day 0. On day 7, cells were dissociated with Accutase and plated at 300,000–450,000 cells/well on Matrigel-coated 6-well plates in 2 mL/well of CDMi supplemented with 5 μM SB, 10 μM CHIR, and 10 μM Y-27632. The medium was changed on days 10 and 12. On day 14, cells were dissociated again with Accutase and seeded at 400,000–800,000 cells/well on Matrigel-coated 6-well plates in 2 mL/well of CDMi supplemented with only 10 μM Y-27632. Three days later (day 17), the medium was switched to S-clone SF-O3 (SF-O3, Sekisui Medical Company, Limited, Japan) supplemented with 0.2–0.4% BSA, 10 ng/mL recombinant human IGF-1 (PeproTech, Inc. NJ, USA), 10 ng/mL recombinant human bFGF (Oriental Yeast Co., Ltd., Japan), 10 ng/mL recombinant human HGF (PeproTech, Inc.), and 0.2 mM 2-mercaptoethanol (2-ME, Nacalai tesque). The medium was then changed twice a week until day 38. Finally, on day 38, the medium was switched to Dulbecco's Modified Eagle Medium (DMEM, Nacalai tesque,

08488-55) supplemented with 2% horse serum (HS, Sigma-Aldrich, H1138), 2 mM L-glutamine (Nacalai tesque), 0.5% Penicillin-Streptomycin Mixed Solution, 5 μM SB, 10 ng/ml recombinant human IGF-1, and 0.1 mM 2-ME. The medium was changed three times per week until FCM analysis on day 82. Cells were cultured on Matrigel or p421[64]. The other differentiation induction methods were the same as described above. FCM analysis was performed on day74–84.

RNA extraction and qRT-PCR

Total RNA was extracted with the ReliaPrep RNA Cell Miniprep system (Promega, WI, USA, Z6012). Complementary DNA was synthesized with ReverTra Ace qPCR RT Kit (TOYOBO Co., Ltd., Japan, FSQ-101). RT-qPCR was carried out with the Power SYBR Green PCR Master Mix (Applied Biosystems, MA, USA) and QuantStudio Real-Time PCR Systems (Thermo Fisher Scientific, Inc.). Technical duplicates were performed for each sample. The primers that were used are listed in **Table 1**.

Gene	Forward Sequence	Reverse Sequence
HPRT1	TGGCGTCGTGATTAGTGATG	ATCTCGAGCAAGACGTTCAG
T	ACCCAGTTCATAGCGGTGAC	CATTGGGAGTACCCAGGTTG
TBX6	AGCCTGTGTCTTTCCATCGT	AGGCTGTCACGGAGATGAAT
SIX1	AGTTCTCGCCTCACAACCAC	ACACCCCTCGACTTCTCCTT
DMRT2	GAACCACCAAGCAAGGACTTC	CCCAGACCCTGAATACTGCAT
PAX3	AGGAAGGAGGCAGAGGAAAG	CAGCTGTTCTGCTGTGAAGG
MYH3	GCAGATTGAGCTGGAAAAGG	TCAGCTGCTCGATCTCTTCA
MYOD1	ACGTGAGGACGAGCATGTG	GTGCAGCGCTTGAGTGTCT
MYOG	TGGGCGTGTAAGGTGTGTAA	CGATGTACTGGATGGCACTG

Table1: Primer List of qRT-PCR

ICC

For correlation analysis between myogenic protein expression on day 38 and MYF5+% on day 82 (Figure 2), cells were fixed in a culture plate with 2% paraformaldehyde (PFA, Nacalai tesque) for 10 min after being washed twice with Dulbecco's Phosphate Buffered Saline (D-PBS, Nacalai tesque). Next, the cells were washed twice with D-PBS and blocked with Blocking One (Nacalai tesque) for 1 hour. Following the blocking process, the cells were stained overnight at 4 °C with the MHC (eBioscience, Inc., CA, USA, 14-6503, 1:800) and MYOD1 (Abcam.plc, UK, ab133627, 1:500) antibodies diluted in D-PBS with 0.2% Triton X-100 (0.2% PBS-T, Sigma-Aldrich) and 10% Blocking One. After washing three times with 0.2% PBS-T, the cells were stained for 1 hour at room temperature with the Alexa Fluor 568 conjugated goat-anti-mouse (Thermo Fisher Scientific, Inc., A11031, 1:500) and Alexa Fluor 488 conjugated goat-anti-rabbit (Thermo Fisher Scientific, Inc., A11034, 1:500) antibodies diluted in 0.2% PBS-T with 10% Blocking One and 1 μg/mL DAPI (Thermo Fisher Scientific, Inc.). Thereafter, the cells were washed two times with 0.2% PBS-T and three times with D-PBS.

For time-course observation of MHC expression, cells were fixed in a culture plate with 4% PFA for 15 min after being washed twice with D-PBS. Next, the cells were washed twice with D-PBS containing 0.1% Triton X-100 (0.1% PBS-T) and blocked with Blocking One for 1 hour. Following the blocking process, the cells were stained overnight at 4 °C with the MHC (1:800) and MYOD1 (1:500) antibodies diluted in 0.1% PBS-T with 5% Blocking One. After washing twice with 0.1% PBS-T, the cells were stained for 1 hour at room temperature with the Alexa Fluor 568 conjugated goat-anti-mouse (1:500) and Alexa Fluor 488 conjugated goat-anti-rabbit (1:500) antibodies diluted in 0.1% PBS-T with 5% Blocking One and 10 μg/mL Hoechst 33342 (Thermo Fisher Scientific, Inc.). Finally, the cells were washed two times with 0.1% PBS-T and three times with D-PBS.

The samples were observed with a BZ-X710 microscope (KEYENCE Corporation, Japan), and the MHC-positive area was calculated using the BZ-X analyzer software (KEYENCE Corporation) or ImageJ.

FCM

The cells were dissociated for 30–60 min at 37 °C with DMEM supplemented with 2 mM L-glutamine, 0.5% Penicillin-Streptomycin Mixed Solution, 10% Dispase II (Godo Shusei Co., Ltd., Japan), 50 μg/mL Collagenase G (Meiji Seika Pharma Co., Ltd., Japan), and 10 μg/mL Collagenase H (Meiji Seika Pharma Co., Ltd.). After incubation, the cells were carefully detached via pipetting and the dissociation solution was removed by centrifugation. Thereafter, the cells were incubated in Accutase for 5 min at 37 °C to ensure only single cells were present. After filtration with a 40 μm cell strainer, followed by washing with Hanks' Balanced Salt Solution with 1% BSA (HBSS, Thermo Fisher Scientific, Inc.), 1000,000 cells were stained with 5 μg/mL Hoechst in HBSS and filtered with a 35 μm cell strainer. The cells were kept on ice until FCM analysis with BD LSRFortessa (Becton, Dickinson and Company, NJ, USA). The final differentiation induction efficiency was assessed using CDH13, a MuSC cell surface marker that I previously identified[46]. FCM was performed following our previously reported protocol[46].

Phase contrast cell imaging

The phase contrast cell images (1920 × 1440 pixels, 8-bit monochrome) were obtained with a BZ-X810 microscope (KEYENCE Corporation) using a 4 × objective lens. The same exposure time and gain value were used in all imaging. Twenty-one images per well were taken at different positions of the center part of the well at eight time points from day 14 to 38. I performed 34 wells of differentiation induction experiments and obtained a total of 5712 images.

Feature extraction

A FFT was applied to each cell image to obtain a FFT spectrum. Next, each component of the FFT spectrum was squared to obtain a FFT power spectrum, and shell integration was applied to the FFT power spectrum to obtain a rotationally symmetric 100-dimensional feature vector (**Figure 3.12**).

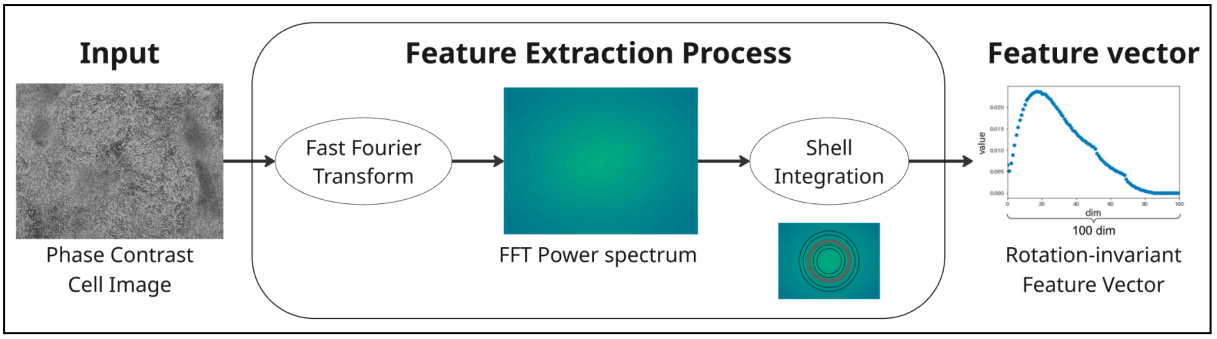


Figure 3.12: Feature Extraction Process overview.

Detailed workflow of the transforming phase contrast cell image into rotation-invariant feature vector.

Classification Tasks

The classifier was trained using the extracted feature vectors and the labels of the images. I labeled each image based on the Task as follows: For Task 1, images with a MYF5+% score of less than 10 were labeled as 1 (positive), and all other images were labeled as 0 (negative). For Task 2, images with a MYF5+% score of more than 20 were labeled as 1 (positive), and all other images were labeled as 0 (negative) (**Figure 3.7A**). In each task, I constructed 50 independent classification models to evaluate accuracy fluctuations due to partitioning by cross-validation. A four-fold cross-validation grouped by well was performed in each classification model to prevent data leakage between training and test sets (**Figure 3.7B**). To generate the input variables for the classifier, I implemented a specific feature representation learning pipeline (**Figure 3.13**). Note that the feature projection models (PCA and PLS) were constructed solely using the training dataset within each cross-validation fold. First, constant variables were removed from the 100-dimensional feature vectors obtained from the FFT. Next, PCA projection was applied to the feature vector. By projecting the vector onto the "Weights of PCA" (top 10 principal components) calculated from the training data, I obtained a compressed 10-dimensional feature vector. Subsequently, PLS projection was performed using this 10-dimensional vector as input. Utilizing the "Weights of PLS" which capture class-discriminative information, I extracted a 1-dimensional feature vector. Finally, these vectors were concatenated to form an 11-dimensional feature vector. This composite vector served as the input for the Random Forest classification model (number of trees = 100). The performance of the trained classifier was assessed by measuring predictive ability on test data not included in the training set, obtaining the average evaluation metrics in validation. Source code is available at (https://github.com/TakuTsuzuki/MuSCs_classifier).

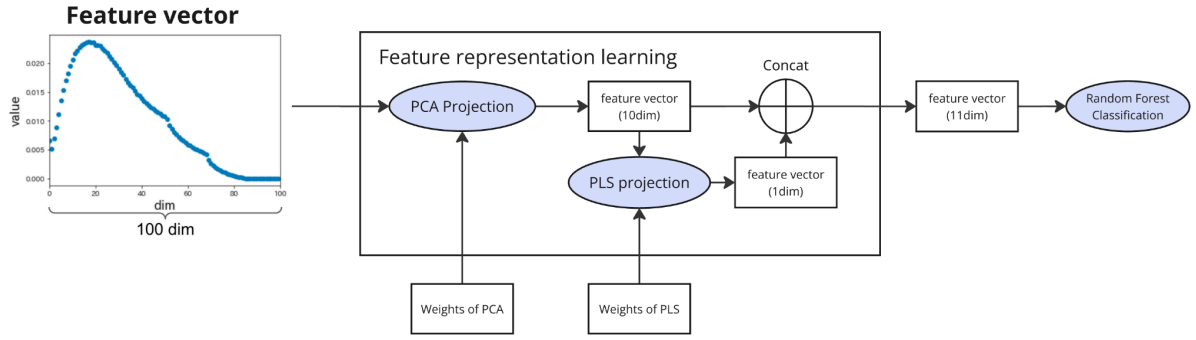


Figure 3.13: Workflow of the dimensionality reduction and feature representation for the Random Forest classifier.

Overview of the feature representation learning pipeline constructed to generate the 11-dimensional input vector for classification. The 100-dimensional feature vector extracted via FFT is first compressed into a 10-dimensional vector using PCA projection. A PLS projection is then applied to the 10-dimensional vector to add a 1-dimensional vector containing class-discriminative information. The final concatenated 11-dimensional vector is used to train the Random Forest classifier. All projection models (PCA and PLS) are constructed solely from the training data to ensure rigorous evaluation.

Statistical analysis

For correlation analysis in **Figures 3.2-3.3**, a test for no correlation was performed. The statistical significance of **Figure 3.5,3.8** was determined using a Mann-Whitney U test. These results were corrected for multiple comparisons using the Benjamini-Hochberg correction.

Ethical approval

Ethical approval for this study was granted by the Graduate School of Medicine, Kyoto University, and Kyoto University Hospital (The approved number R91).

Data availability

Source code that supports the results is available in (https://github.com/TakuTsuzuki/MuSCs_classifier). The data will be made available by the corresponding author upon reasonable request.

References

1. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).

2. Yamanaka, S. Pluripotent Stem Cell-Based Cell Therapy—Promise and Challenges.

Cell Stem Cell **27**, 523–531 (2020).

3. Liu, G., David, B. T., Trawczynski, M. & Fessler, R. G. Advances in Pluripotent Stem Cells: History, Mechanisms, Technologies, and Applications. *Stem Cell Rev Rep* **16**, 3–32 (2020).
4. Ohnuki, M. & Takahashi, K. Present and future challenges of induced pluripotent stem cells. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20140367 (2015).
5. Jin, Z.-B. *et al.* Modeling retinal degeneration using patient-specific induced pluripotent stem cells. *PLoS One* **6**, e17084 (2011).
6. Mandai, M. *et al.* Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. *N. Engl. J. Med.* **376**, 1038–1046 (2017).
7. Chal, J. & Pourquié, O. Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **144**, 2104–2122 (2017).
8. Otsuka, Y. *et al.* One-step induction of photoreceptor-like cells from human iPSCs by delivering transcription factors. *iScience* **25**, 103987 (2022).
9. Tanaka, A. *et al.* Efficient and reproducible myogenic differentiation from human iPS cells: prospects for modeling Miyoshi Myopathy in vitro. *PLoS One* **8**, e61540 (2013).
10. Zhao, M. *et al.* Induced Fetal Human Muscle Stem Cells with High Therapeutic Potential in a Mouse Muscular Dystrophy Model. *Stem Cell Reports* **15**, 80–94 (2020).
11. Shoji, E. *et al.* Early pathogenesis of Duchenne muscular dystrophy modelled in patient-derived human induced pluripotent stem cells. *Sci. Rep.* **5**, 12831 (2015).
12. Song, B. *et al.* The directed differentiation of human iPS cells into kidney podocytes. *PLoS One* **7**, e46453 (2012).
13. Nakajima, T. *et al.* Grafting of iPS cell-derived tenocytes promotes motor function recovery after Achilles tendon rupture. *Nat. Commun.* **12**, 5012 (2021).
14. Chal, J. *et al.* Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat. Biotechnol.* **33**, 962–969 (2015).
15. Tsujimoto, H. *et al.* A Modular Differentiation System Maps Multiple Human Kidney Lineages from Pluripotent Stem Cells. *Cell Rep.* **31**, 107476 (2020).

16. Yamashita, A. *et al.* Generation of scaffoldless hyaline cartilaginous tissue from human iPSCs. *Stem Cell Reports* **4**, 404–418 (2015).
17. Kikuchi, T. *et al.* Human iPS cell-derived dopaminergic neurons function in a primate Parkinson's disease model. *Nature* vol. 548 592–596 Preprint at https://doi.org/10.1038/nature23664 (2017).
18. Kamao, H. *et al.* Characterization of human induced pluripotent stem cell-derived retinal pigment epithelium cell sheets aiming for clinical application. *Stem Cell Reports* **2**, 205–218 (2014).
19. Fujimori, K. *et al.* Escape from Pluripotency via Inhibition of TGF-β/BMP and Activation of Wnt Signaling Accelerates Differentiation and Aging in hPSC Progeny Cells. *Stem Cell Reports* **9**, 1675–1691 (2017).
20. Ma, X. *et al.* Highly efficient differentiation of functional hepatocytes from human induced pluripotent stem cells. *Stem Cells Transl. Med.* **2**, 409–419 (2013).
21. Numasawa-Kuroiwa, Y. *et al.* Involvement of ER stress in dysmyelination of Pelizaeus-Merzbacher Disease with PLP1 missense mutations shown by iPSC-derived oligodendrocytes. *Stem Cell Reports* **2**, 648–661 (2014).
22. Hayashi, Y., Ohnuma, K. & Furue, M. K. Pluripotent Stem Cell Heterogeneity. *Adv. Exp. Med. Biol.* **1123**, 71–94 (2019).
23. Kajiwara, M. *et al.* Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12538–12543 (2012).
24. Koyanagi-Aoi, M. *et al.* Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20569–20574 (2013).
25. Nishizawa, M. *et al.* Epigenetic Variation between Human Induced Pluripotent Stem Cell Lines Is an Indicator of Differentiation Capacity. *Cell Stem Cell* **19**, 341–354 (2016).
26. Nayak, P., Colas, A., Mercola, M., Varghese, S. & Subramaniam, S. Temporal mechanisms of myogenic specification in human induced pluripotent stem cells. *Sci Adv*

7, (2021).

27. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* **28**, 848–855 (2010).
28. Liang, G. & Zhang, Y. Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* **13**, 149–159 (2013).
29. Tristan, C. A. *et al.* Robotic high-throughput biomanufacturing and functional differentiation of human pluripotent stem cells. *Stem Cell Reports* **16**, 3076–3092 (2021).
30. Kanie, K. *et al.* Effect of mechanical vibration stress in cell culture on human induced pluripotent stem cells. *Regen Ther* **12**, 27–35 (2019).
31. Cerneckis, J., Cai, H. & Shi, Y. Induced pluripotent stem cells (iPSCs): molecular mechanisms of induction and applications. *Signal Transduct. Target. Ther.* **9**, 112 (2024).
32. Nakajima, T. *et al.* Modeling human somite development and fibrodysplasia ossificans progressiva with induced pluripotent stem cells. *Development* **145**, (2018).
33. Sakurai, H. *et al.* Bidirectional induction toward paraxial mesodermal derivatives from mouse ES cells in chemically defined medium. *Stem Cell Res.* **3**, 157–169 (2009).
34. Park, R. *et al.* Regulated behavior in living cells with highly aligned configurations on nanowrinkled graphene oxide substrates: Deep learning based on interplay of cellular contact guidance. *ACS Nano* **18**, 1325–1344 (2024).
35. Kusumoto, D. *et al.* Automated Deep Learning-Based System to Identify Endothelial Cells Derived from Induced Pluripotent Stem Cells. *Stem Cell Reports* **10**, 1687–1695 (2018).
36. Niioka, H. *et al.* Classification of C2C12 cells at differentiation by convolutional neural network of deep learning using phase contrast images. *Hum. Cell* **31**, 87–93 (2018).
37. Schaub, N. J. *et al.* Deep learning predicts function of live retinal pigment epithelium from quantitative microscopy. *J. Clin. Invest.* **130**, 1010–1023 (2020).
38. Orita, K., Sawada, K., Koyama, R. & Ikegaya, Y. Deep learning-based quality control of

cultured human-induced pluripotent stem cell-derived cardiomyocytes. *J. Pharmacol. Sci.* **140**, 313–316 (2019).

39. Imamura, K. *et al.* Prediction Model of Amyotrophic Lateral Sclerosis by Deep Learning with Patient Induced Pluripotent Stem Cells. *Ann. Neurol.* **89**, 1226–1233 (2021).
40. Heigwer, F. *et al.* A global genetic interaction network by single-cell imaging and machine learning. *Cell Syst.* **14**, 346–362.e6 (2023).
41. Shakarchy, A. *et al.* Machine learning inference of continuous single-cell state transitions during myoblast differentiation and fusion. *Mol. Syst. Biol.* **20**, 217–241 (2024).
42. Tohgasaki, T., Touyama, A., Kousai, S. & Imai, K. Machine learning-enhanced estimation of cellular protein levels from bright-field images. *Bioengineering (Basel)* **11**, 774 (2024).
43. Kazemimoghadam, M., Janmaleki, M., Fouani, M. H. & Abbasi, S. Prediction of neural differentiation fate of rat mesenchymal stem cells by quantitative morphological analyses using image processing techniques. *Biomed. Tech.* **60**, 57–64 (2015).
44. Buggenthin, F. *et al.* Prospective identification of hematopoietic lineage choice by deep learning. *Nat. Methods* **14**, 403–406 (2017).
45. Matsuoka, F. *et al.* Morphology-based prediction of osteogenic differentiation potential of human mesenchymal stem cells. *PLoS One* **8**, e55082 (2013).
46. Nalbandian, M. *et al.* Characterization of hiPSC-Derived Muscle Progenitors Reveals Distinctive Markers for Myogenic Cell Purification Toward Cell Therapy. *Stem Cell Reports* **16**, 883–898 (2021).
47. Beauchamp, J. R. *et al.* Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J. Cell Biol.* **151**, 1221–1234 (2000).
48. Hashimoto, K., Fujimoto, H. & Nakatsuji, N. An ECM substratum allows mouse mesodermal cells isolated from the primitive streak to exhibit motility similar to that inside the embryo and reveals a deficiency in the T/T mutant cells. *Development* vol. 100 587–598 Preprint at https://doi.org/10.1242/dev.100.4.587 (1987).

49. Chapman, D. L., Agulnik, I., Hancock, S., Silver, L. M. & Papaioannou, V. E. Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Dev. Biol.* **180**, 534–542 (1996).
50. Nord, H., Nygård Skalman, L. & von Hofsten, J. Six1 regulates proliferation of Pax7-positive muscle progenitors in zebrafish. *J. Cell Sci.* **126**, 1868–1880 (2013).
51. Sato, T., Rocancourt, D., Marques, L., Thorsteinsdóttir, S. & Buckingham, M. A Pax3/Dmrt2/Myf5 regulatory cascade functions at the onset of myogenesis. *PLoS Genet.* **6**, e1000897 (2010).
52. Karsch-Mizrachi, I., Travis, M., Blau, H. & Leinwand, L. A. Expression and DNA sequence analysis of a human embryonic skeletal muscle myosin heavy chain gene. *Nucleic Acids Res.* **17**, 6167–6179 (1989).
53. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
54. Wright, W. E., Sassoon, D. A. & Lin, V. K. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* **56**, 607–617 (1989).
55. Shahinfar, S., Meek, P. & Falzon, G. 'How many images do I need?' Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecol. Inform.* **57**, 101085 (2020).
56. Ruggeri, A., Grisan, E. & Jaroszewski, J. A new system for the automatic estimation of endothelial cell density in donor corneas. *Br. J. Ophthalmol.* **89**, 306–311 (2005).
57. Ho, T. K. Random decision forests. in *Proceedings of 3rd International Conference on Document Analysis and Recognition* vol. 1 278–282 vol.1 (1995).
58. Bradley, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **30**, 1145–1159 (1997).
59. Bischoff, R. Interaction between satellite cells and skeletal muscle fibers. *Development* **109**, 943–952 (1990).
60. Nakagawa, M. *et al.* A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. *Sci. Rep.* **4**, 3594 (2014).

61. Kawada, R. *et al.* Establishment of quantitative and consistent in vitro skeletal muscle pathological models of myotonic dystrophy type 1 using patient-derived iPSCs. *Sci. Rep.* **13**, 94 (2023).
62. Sasaki-Honda, M. *et al.* A patient-derived iPSC model revealed oxidative stress increases facioscapulohumeral muscular dystrophy-causative DUX4. *Hum. Mol. Genet.* **27**, 4024–4035 (2018).
63. Wada, E. *et al.* Optimized simple culture protocol for inducing mature myotubes from MYOD1-overexpressed human iPS cells. *Sci. Rep.* **14**, 28783 (2024).
64. Zhao, M. *et al.* Heparan sulfate chain-conjugated laminin-E8 fragments advance paraxial mesodermal differentiation followed by high myogenic induction from hiPSCs. *Adv. Sci. (Weinh.)* **11**, e2308306 (2024).

Chapter4

4. General Discussion and Conclusion

4.1 Chapter Summary

Chapter 4 synthesizes the contributions of this thesis under the Closed-Loop Scientific Discovery System (CL-SDS) framework introduced in Chapter 1. Section 4.2 first summarizes the three primary contributions of this dissertation. The first contribution identifies the domain of regenerative medicine process development as a critical "white space" characterized by the triple barriers of high dimensionality, physical sensitivity, and time latency. This domain represents a setting where the strategic value of CL-SDS is maximized (Chapter 1). The second contribution represents the first physical instantiation of CL-SDS in this domain, achieving autonomous knowledge discovery. By integrating the humanoid robot Maholo, batch Bayesian optimization, and automated image scoring, I constructed a system that navigated a seven-variable search space. Through three iterative rounds of 48-parallel experiments, the system discovered a protocol that improved hiPSC-to-RPE differentiation efficiency (pigmentation score) by 88% relative to the pre-optimized baseline (Chapter 2). The third contribution addresses the barrier of time latency in long-horizon protocols. I demonstrated that terminal quality can be predicted more than 50 days in advance using phase-contrast imaging. Specifically, in a hiPSC-to-MuSC protocol, the differentiation index MYF5+% at day 82 was successfully predicted from images acquired during the myogenic induction phase (days 24–31) (Chapter 3). In Section 4.3, I integrate these findings to propose a next-generation system architecture. Building on the two system primitives established in this thesis, specifically (i) the physically closed loop (Chapter 2) and (ii) intermediate-state observability via early prediction (Chapter 3), I present a unified vision for advancing process development in two directions: faster discovery and robust manufacturing. Specifically, I discuss accelerated discovery through multi-fidelity (grey-box) closed-loop optimization, which utilizes intermediate predictions as low-cost “weak observations” to shorten feedback cycles, and adaptive manufacturing through trajectory-level feedback control, which modifies control inputs based on estimated states to absorb intrinsic biological variability.

4.2 Summary of Contributions

This dissertation advances the development of Closed-Loop Scientific Discovery Systems (CL-SDS) for regenerative medicine process development by (i) identifying this domain as a critical “white space” characterized by specific barriers and formalizing it as a search-and-control problem, (ii) physically closing the loop with an AI-robotic optimizer to demonstrate autonomous protocol improvement despite high dimensionality and physical sensitivity, and (iii) introducing intermediate-state observability via early, non-destructive outcome prediction to mitigate the time-latency bottleneck in long-horizon differentiation protocols. Together, these contributions establish two system primitives, specifically physical loop closure and intermediate-state observability, which enable the accelerated and adaptive CL-SDS architectures discussed in Section 4.3.

Contribution1: Framing regenerative medicine process development as a high-impact “white space” for CL-SDS, and formalizing the system definition (Chapter 1)

Chapter 1 establishes CL-SDS as a cyber-physical architecture in which a Problem Definition (Search Space X, Objective y, Budgets B) is executed by a Search Strategy P through an autonomous loop of Model Generation, Experimental Design, and Physical Execution. This formalization clarifies what must be specified externally versus what should be executed autonomously, thereby providing a precise language for system design, comparison, and extension. Crucially, Chapter 1 positions regenerative-medicine manufacturing process development as an "unexplored frontier" (white space) where the strategic value of CL-SDS is maximized, grounded in the principle that "the process is the product." This frontier is characterized by a triple barrier consisting of High Dimensionality, Physical Sensitivity, and Time Latency. This combination makes conventional human-centered process development structurally brittle and difficult to reproduce. The chapter further articulates a thesis-level mapping from these barriers to the subsequent technical chapters as a coherent, system-level research program.

Contribution2: Physical loop closure for high-dimensional, physically sensitive protocol optimization via CL-SDS (Chapter 2)

Chapter 2 provides the first instantiation of CL-SDS for regenerative-medicine process development by integrating laboratory robotics, batch Bayesian optimization, and automated image scoring into an end-to-end closed loop for hiPSC differentiation. The system operationalizes induced differentiation as a black-box optimization problem whose learnability depends critically on suppressing run-to-run variability at the data-generation layer. This is precisely the regime where robotic reproducibility and AI-driven planning are complementary. Concretely, the hiPSC to RPE differentiation protocol was parameterized as a seven-variable search space, and the CL-SDS executed three iterative rounds of 48-well parallel experiments. Through this autonomous, reproducible exploration, the system discovered a protocol that improved hiPSC-to-RPE differentiation efficiency by 88% relative to the pre-optimized baseline, as quantified by the pigmentation score. Beyond the performance gain itself, this result demonstrates that a historically tacit, skill-dependent practice can be reframed as an executable closed-loop optimization pipeline in which

robotics reduces physical sensitivity and AI systematically navigates high-dimensional design spaces.

Contribution3: Time-axis compression via early, non-destructive prediction of terminal outcomes from intermediate images (Chapter 3)

Chapter 3 addresses the remaining bottleneck of Time Latency. This challenge persists even after physical loop closure because many regenerative protocols require months before terminal phenotypes become measurable. Consequently, in a sequential optimization loop, examining a sufficient number of experimental conditions becomes prohibitively slow due to the extended feedback cycle inherent to long-term culture. To mitigate this rate limitation, Chapter 3 establishes that long-horizon differentiation trajectories contain informative, non-destructive signatures at intermediate stages that can be leveraged as an early readout. Specifically, in a hiPSC to MuSC differentiation protocol, the study demonstrates that the terminal endpoint quality (specifically MYF5+% on day 82) can be predicted more than 50 days in advance from phase-contrast images acquired during the induction phase (days 14 to 38). Informative time windows are already apparent in the early induction period, such as around days 24 to 30. This contribution provides a concrete foundation for shortening the effective feedback timescale in process development by enabling earlier decision-making, such as early filtering, termination, or surrogate evaluation, without destructive assays.

Taken together, Chapters 2 and 3 establish two complementary system primitives for regenerative-medicine CL-SDS: (i) physical loop closure through AI-robotic execution and (ii) intermediate-state observability through early prediction. Section 4.3 builds directly on this combination to discuss how integrating the CL-SDS from Chapter 2 with the early prediction capabilities of Chapter 3 (intermediate state estimation) enables regenerative-medicine process development to proceed in two distinct ways. First, it allows for faster process development through accelerated discovery via multi-fidelity closed-loop optimization. Second, it enables stronger results through robust manufacturing via trajectory-level feedback control and adaptive culture operation.

4.3 Discussion: Toward Accelerated & Adaptive CL-SDS

Chapters 2 and 3 established two complementary “system primitives” for Closed-Loop Scientific Discovery Systems (CL-SDS) in regenerative medicine process development: (i) AI-robotic protocol-optimization loop that can reproducibly execute and optimize high-dimensional cell-culture procedures (Chapter 2), and (ii) an early, non-destructive prediction module that can infer terminal differentiation outcomes from intermediate morphological information (Chapter 3). Together, these results suggest that CL-SDS can be extended from “optimize a static protocol and wait for the terminal readout” toward a system that actively exploits intermediate information to reduce feedback latency and, ultimately, to support adaptive intervention during culture.

This integration naturally connects our platform to two adjacent bodies of literature. First, from an optimization perspective, early prediction can be interpreted as a low-fidelity evaluation of the final objective, which reframes long-horizon cell culture as a multi-fidelity

optimization problem—where information with different cost/accuracy levels is orchestrated to accelerate learning and convergence [1–3]. Second, from a process-systems perspective, intermediate prediction can be interpreted as a form of state estimation, which is a prerequisite for modern feedback control and quality-assurance frameworks in bioprocessing (12,13). In the remainder of Section 4.3, I focus first on the acceleration mechanism: how early prediction changes the "clock speed" of CL-SDS by shortening the effective feedback cycle (Section 4.3.1). I then position the same observability as an enabling layer for adaptive, trajectory-level process control (Section 4.3.2).

4.3.1 Accelerated discovery through multi-fidelity closed-loop optimization

(1) Current Challenge

Even after a robotic CL-SDS is physically realized (Chapter 2), regenerative medicine process development remains constrained by a fundamental property of living systems: the objective of interest is often observable only after a long culture period. In other words, each "query" (a protocol condition) is an expensive, high-latency evaluation that may take months before yielding the terminal readout. This latency is not a statement that the experiments are infeasible but it implies that, under a fixed wall-clock research horizon, the number of decision points (optimization rounds) is severely limited.

This limitation matters because the underlying search problem is typically high-dimensional and non-convex, where performance improvements accumulate through iterative model updates and sequential experimentation. In a naive CL-SDS that relies only on terminal readouts, the loop can rotate only after the full culture period completes. Conceptually (**Figure. 4.1A**), the system performs Round 1, waits until the end to obtain high-fidelity scores, and only then designs Round 2. The consequence is a slow "outer loop": for example, if one full evaluation effectively consumes ~3 months, only ~2 rounds can be executed per half-year—often insufficient to reliably approach a high-performing region in a large search space. A second, subtler inefficiency is that **the culture trajectory contains information long before the terminal assay**, yet a naive loop discards this intermediate signal and uses only the final endpoint value for learning. In long-horizon biology, this means the system is not fully utilizing the information content generated during the waiting period.

(2) Solution

The key idea is to treat early prediction (Chapter 3) as a low-fidelity proxy for the terminal objective and inject it into the CL-SDS loop as soon as intermediate observations become available. This changes the planning problem from "single-fidelity, end-point optimization" into a multi-fidelity optimization problem [1–3]:

- **High fidelity:** the terminal readout (ground truth) obtained after completing the full culture period.
- **Low fidelity:** an early prediction of that terminal readout computed from intermediate, non-destructive observations (e.g., imaging-based morphological signatures).

Under this view, the system can update its surrogate model and propose subsequent experiments using the low-fidelity signal, while still reserving high-fidelity evaluations to verify and calibrate the search. In practice, the integration can be implemented in an

asynchronous/parallel Bayesian optimization style [4]: as soon as low-fidelity predictions are produced, the optimizer can generate the next batch of candidate conditions without waiting for all high-fidelity results to complete. This is conceptually illustrated in **Figure. 4.1B**: early prediction produces interim scores during Round 1, enabling Round 2 planning to start earlier; repeating this mechanism allows an additional optimization round (Round 3) to be executed within the same wall-clock horizon.

Importantly, the role of early prediction here is not to "replace" the ground truth. Rather, it provides a **time-advantaged signal** that enables earlier decision-making. The system can then employ a principled allocation strategy—analogous to resource-allocation ideas in multi-fidelity optimization and early-stopping literature —in which Low-fidelity evaluations are used broadly to explore and rapidly filter/triage conditions, and high-fidelity evaluations are used selectively to confirm the most promising candidates and to prevent optimization drift caused by proxy bias.

In this section, I intentionally focus on the time-latency benefit under the simplifying assumption that sufficient physical capacity exists to overlap experiments; the central claim is that early prediction reduces the effective feedback delay and increases the frequency of model updates and experimental design decisions.

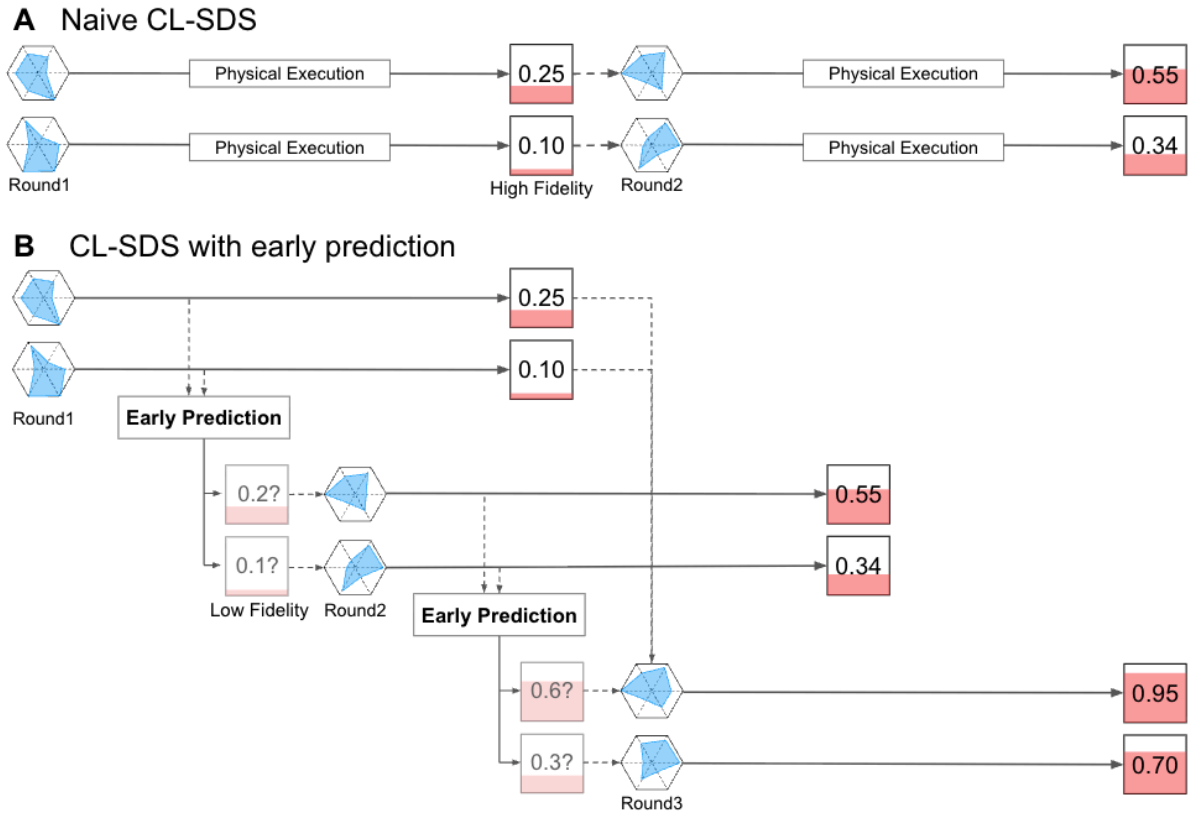


Fig 4.1: Conceptual acceleration of CL-SDS by injecting early outcome prediction as a low-fidelity signal.

(A) Naive CL-SDS (single-fidelity loop). Each round consists of physical execution (cell culture under a candidate condition) followed by a high-fidelity terminal evaluation available only after completing the full culture period. The optimizer updates its model and proposes the next round only after the terminal scores are obtained, limiting the number of rounds achievable under a fixed wall-clock budget.

(B) CL-SDS with early prediction (multi-fidelity loop). During culture, an early prediction module produces low-fidelity estimates of the terminal outcome from intermediate

observations. These low-fidelity signals are fed back to the optimizer to enable earlier planning of subsequent rounds (dashed arrows), allowing additional optimization cycles within the same wall-clock time. Terminal high-fidelity evaluations remain the ground truth used for confirmation and calibration. Numeric boxes represent illustrative objective values; lighter boxes indicate low-fidelity predicted scores, and darker boxes indicate high-fidelity terminal scores.

(3) Expected Effect

The integration of early prediction compresses the effective feedback time of the CL-SDS by decoupling the decision cycle from the terminal biological readout. Rather than waiting until Day T to obtain the final score for every candidate, the system can make earlier go/no-go decisions and update its model at an intermediate time point ($t \ll T$), then immediately launch the next experimental round (**Figure. 4.1**). As a result, within a fixed wall-clock horizon H, the number of sequential learning cycles increases from approximately H/T to H/t. Because each cycle refines the experiment-design policy in a high-dimensional search, this increase in sequential updates is the direct mechanism through which the system attains higher-performing protocols within the same total research period.

(4) Risks

Integrating early prediction shortens feedback latency but introduces three specific failure modes. First, proxy bias (fidelity gap) occurs if the early signal misrepresents terminal quality, causing the optimizer to confidently converge to the wrong region and discard promising conditions via misranking. Second, instability under distribution shift arises when predictions are made outside the training distribution; to prevent noise from derailing the search, the system requires uncertainty-aware gating that routes low-confidence candidates to high-fidelity validation. Third, the resource-allocation dilemma presents a trade-off where over-using expensive assays negates acceleration while over-using cheap proxies increases error; thus, the optimal fidelity balance must be calibrated through iterative deployment rather than fixed a priori.

4.3.2 Reducing quality variability via state estimation & feedback control (Especially autologous products)

(1) Current Challenge

A central bottleneck in regenerative-medicine manufacturing is not only improving the average outcome, but reducing outcome variability under inevitable input uncertainty. In autologous and patient-specific settings, the initial biological state of the starting material (e.g., cell fitness, activation history, epigenetic status) can differ substantially across patients and lots, even when the downstream protocol and equipment are robotized and standardized. This “initial-condition uncertainty” propagates along a long culture trajectory and appears as batch-to-batch fluctuation in final CQAs (or outright batch failure), which directly undermines process capability and supply reliability for cell-based therapies [5,6].

Conventional manufacturing relies on a static (open-loop) protocol: the same schedule of feeds, medium exchanges, and handling inputs is applied to all lots. As conceptualized in **Figure. 4.2A**, even if upstream variability is unavoidable, an open-loop protocol has no mechanism to (i) detect whether an individual trajectory is drifting toward an

off-spec region, nor (ii) correct that drift with adaptive interventions. As a result, variability at input is mapped into variability at output, yielding inconsistent final quality [5].

(2) Solution

The primary solution is adaptive manufacturing: using intermediate, non-destructive signals as process observations to drive feedback control that actively suppresses quality dispersion (**Figure. 4.2B**). This structure is naturally decomposed into two functional requirements—**state estimation** and **control policy**—and clarifies where CL-SDS contributes.

Step 1 — State estimation (observability): "What state is this culture in right now?"

Adaptive control requires that the latent biological trajectory be "made observable" through measurable proxies. In bioprocess engineering, this corresponds to soft sensing / state estimation, where partial measurements are converted into an estimate of internal process state suitable for decision-making [7–9].

In this thesis, Chapter 3 establishes a concrete state-estimation primitive: a non-destructive imaging-based predictor that maps intermediate morphology to future terminal outcomes, demonstrating that long-horizon differentiation trajectories contain early signatures informative of final fate. In the adaptive-manufacturing framing, this predictor functions as an on-line estimator of "where the culture is headed," providing a practical intermediate state variable for closed-loop control.

Step 2 — Feedback control policy (controllability): "What input should be applied to reduce variance?"

Once state is estimated, the next problem is to define and implement an intervention rule (a control policy) that selects actionable inputs—e.g., timing/intensity of medium exchange, passaging schedule, cytokine/chemical dosing schedule, environmental setpoints, or other protocol levers—to steer trajectories back toward a target quality region. This is the core mechanism depicted in **Figure. 4.2B**: the controller uses the estimated state to choose interventions that "pull" trajectories inward, narrowing the output distribution.

A practical precedent of this principle already exists in cell culture automation: Ochiai et al. implemented variable scheduling based on measured culture progression, autonomously adjusting operation timing to stabilize culture conditions [10]. The present direction generalizes this idea from "schedule stabilization" to quality stabilization, where the feedback variable is an intermediate predictor linked to terminal CQAs rather than a single operational proxy.

Where CL-SDS becomes essential ?: learning the control policy under biological uncertainty

Crucially, the hard part is typically not writing down the feedback concept, but determining which interventions reduce variance, under which estimated states, and with what trade-offs (risk, cost, time). In other words, the controller itself must be designed/learned as a data-driven object. This is where CL-SDS plays a principled role—not by replacing feedback control, but by providing an efficient experimental engine to learn the policy. Concretely, policy learning can be formulated as a sequential decision/optimization problem over repeated cultures:

- **Design-time loop (across runs):** use an optimizer (e.g., Bayesian optimization) to explore candidate policy parameters; execute cultures; quantify robustness (e.g., probability of meeting CQA, variance reduction); update the model; propose improved policies. This is a CL-SDS loop whose "search space" is the controller/policy class rather than static protocol parameters [11].
- **Run-time loop (within a run):** use the learned policy to adapt inputs based on the estimated intermediate state (**Figure. 4.2B**). This is the feedback-control loop, grounded in state estimation and control [8,9].

Depending on how the controller is represented, this learning can be model-based (e.g., system identification + model predictive control) [9,12] or model-free (e.g., reinforcement-learning-style policy improvement) [13–15]. The key point for this thesis is architectural: Chapter 3 provides the observability primitive, and a CL-SDS framework provides the systematic way to learn the intervention rule needed to achieve robustness against patient/lot variability.

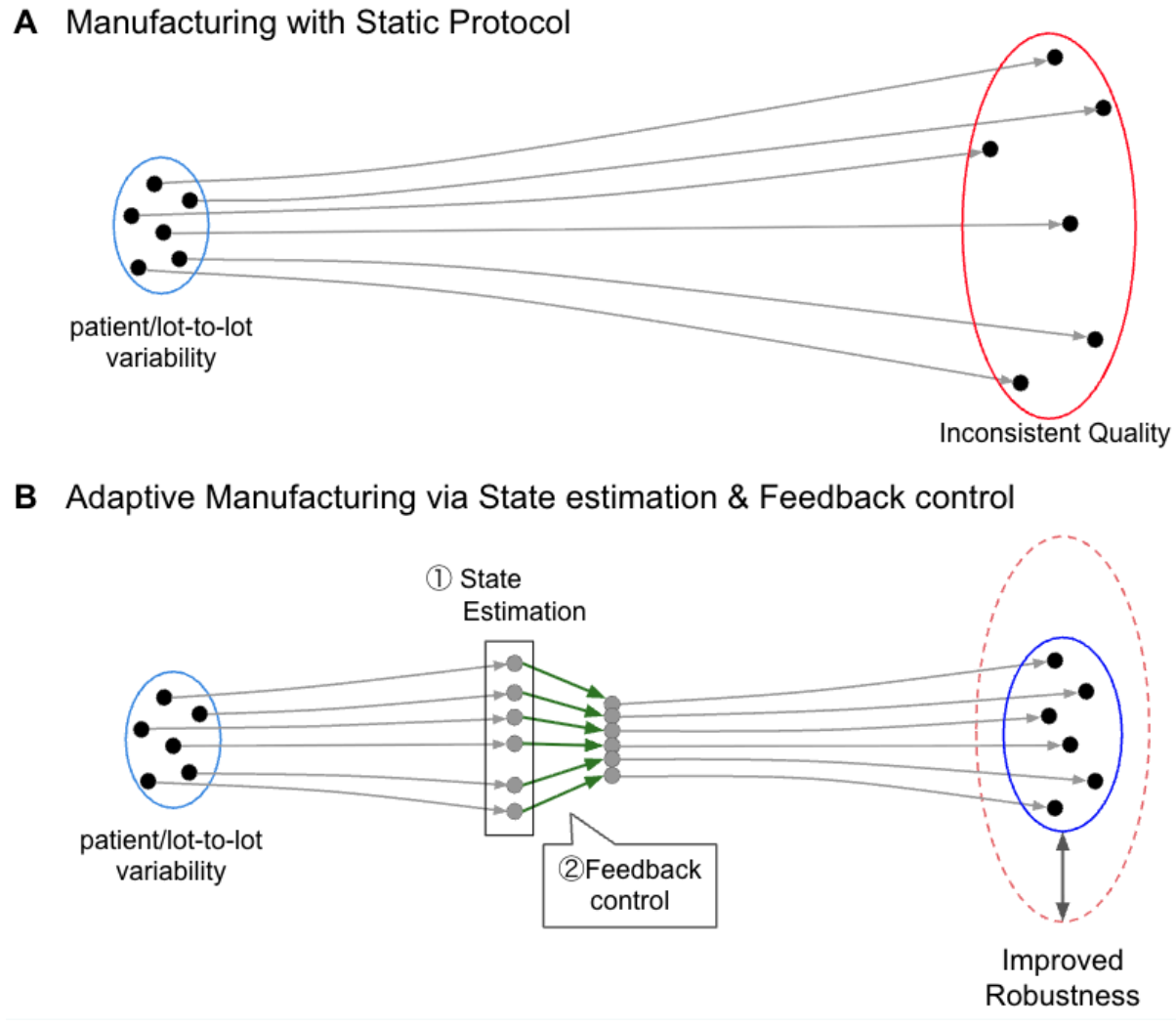


Figure 4.2: From static protocol manufacturing to adaptive manufacturing enabled by state estimation and feedback control.
(A) Manufacturing with a static protocol. Patient/lot-to-lot variability in starting material propagates through a fixed, static protocol, producing divergent culture trajectories and a wide distribution of terminal quality (inconsistent outcome quality).

(B) Adaptive manufacturing via state estimation and feedback control. Intermediate, non-destructive observations are used by an observer (state estimator) to infer the latent culture state during execution. A controller then adaptively modulates controllable process inputs (e.g., scheduling and dosing decisions) to steer trajectories toward a reference region associated with acceptable quality, reducing outcome variance and improving robustness.

(3) Expected Effect

By closing the loop at the trajectory level, adaptive manufacturing transforms unavoidable input variability into a controllable disturbance: intermediate-state estimation enables early detection of drift, and feedback control applies corrective inputs to keep cultures within a target quality envelope. The expected outcome is improved robustness—i.e., a narrower distribution of final quality and a higher probability of meeting CQAs across heterogeneous lots—thereby reducing batch failure risk and improving manufacturing reliability (**Figure. 4.2B**).

(4) Risks

Trajectory-level feedback control is powerful but subject to four fundamental constraints. First, observability limits imply that if observations fail to encode the latent state, identical inputs will yield divergent outcomes, preventing policy convergence. Second, controllability and action-space design dictate that poor variable selection can render the system observable but practically uncontrollable. Third, safety and stability are critical because closing the loop can unintentionally amplify noise or degrade terminal quality for short-term gains; therefore, interventions require strict constraints and fail-safe mechanisms that revert to a baseline upon anomaly detection. Fourth, non-stationarity due to biological drift means a fixed controller will degrade over time, necessitating lifecycle strategies like distribution-shift detection and periodic re-calibration.

4.4 Conclusion

This thesis set out to operationalize a **Closed-Loop Scientific Discovery System (CL-SDS)** for manufacturing process development in regenerative medicine—an experimental domain that is simultaneously constrained by (i) combinatorially vast search spaces, (ii) extreme sensitivity to physical manipulation, and (iii) long feedback latency inherent to differentiation protocols. By treating process development not as an artisanal activity driven by tacit skills but as an algorithmic search problem that can be executed and iterated by machines, this thesis advances a concrete pathway to transform regenerative medicine process development into a computable engineering discipline.

A key conceptual contribution of this work is the disciplined framing of discovery and process development through CL-SDS: a cyber-physical loop in which a user-specified problem definition and search strategy drive iterative cycles of model construction, experimental design, and physical execution. This framing is particularly consequential for regenerative medicine, where the principle that “the process is the product” makes process knowledge inseparable from product identity and quality. In this context, the central challenge is not simply to automate predetermined protocols, but to build systems that can systematically explore, evaluate, and improve the process itself under realistic experimental budgets and noise.

Chapter 2 establishes the first instantiation of CL-SDS for regenerative medicine process development. By integrating laboratory robotics, batch Bayesian optimization, and automated image scoring, the system operationalized induced differentiation as a black-box optimization problem where robotic reproducibility was critical for suppressing run-to-run variability. Specifically, the hiPSC-to-RPE differentiation protocol was parameterized as a seven-variable search space, and the system executed three iterative rounds of 48-well parallel experiments. Through this autonomous approach, the CL-SDS discovered a condition that improved differentiation efficiency by 88% relative to the baseline condition based on the pigmentation score. Importantly, biological validation confirmed that the robot-manufactured cells satisfied key characteristics for regenerative medicine, demonstrating that a historically tacit, skill-dependent practice can be successfully reframed as an executable closed-loop optimization pipeline.

Chapter 3 addressed a bottleneck that persists even after cyber-physical loop closure: the time cost imposed by long differentiation trajectories. I developed an early, non-destructive prediction method that infers terminal induction efficiency from intermediate morphology captured by phase-contrast imaging. In a directed differentiation protocol from hiPSCs to muscle stem cells (MuSc), the terminal endpoint (MYF5+% on day 82) was shown to be predictable from intermediate images acquired during the myogenic induction phase. Using a robust classical pipeline designed for small datasets, the approach demonstrated stable predictive performance (on the order of ~0.8 AUC) and enabled discrimination of high- versus low-efficiency outcomes approximately 50 days before the endpoint. This result establishes that long-horizon differentiation trajectories contain early morphological signatures correlated with terminal fate, and that such signatures can be exploited to shorten the effective feedback cycle without destructive assays.

Taken together, this dissertation contributes two system primitives required to make CL-SDS practical for regenerative medicine process development: (1) reproducible and scalable cyber-physical execution that supports sample-efficient optimization in high-dimensional spaces, and (2) intermediate-state observability via early, non-destructive prediction that reduces the wall-clock cost of learning. By demonstrating both primitives in biologically and operationally realistic case studies, this work advances CL-SDS from a conceptual architecture to an actionable methodology for accelerating, standardizing, and systematically improving cell-manufacturing processes in regenerative medicine.

References

1. Kennedy, M. Predicting the output from a complex computer code when fast approximations are available. *Biometrika* **87**, 1–13 (2000).
2. Forrester, A. I. J., Sóbester, A. & Keane, A. J. Multi-fidelity optimization via surrogate modelling. *Proc. Math. Phys. Eng. Sci.* **463**, 3251–3269 (2007).
3. Kandasamy, K., Dasarathy, G., Schneider, J. & Poczos, B. Multi-fidelity Bayesian optimisation with continuous approximations. *arXiv [stat.ML]* 1799–1808 (2017).
4. Kandasamy, K., Krishnamurthy, A., Schneider, J. & Poczos, B. Asynchronous parallel

Bayesian optimisation via Thompson sampling. *arXiv [stat.ML]* (2017).

5. Lipsitz, Y. Y., Timmins, N. E. & Zandstra, P. W. Quality cell therapy manufacturing by design. *Nat. Biotechnol.* **34**, 393–400 (2016).
6. Ayala Ceja, M., Khericha, M., Harris, C. M., Puig-Saus, C. & Chen, Y. Y. CAR-T cell manufacturing: Major process parameters and next-generation strategies. *J. Exp. Med.* **221**, (2024).
7. Quality Assurance. Guidance for Industry. https://www.fda.gov/media/71012/download.
8. Bastin, G. & Dochain, D. *On-Line Estimation and Adaptive Control of Bioreactors*. (Elsevier Science & Technology, 2013).
9. Kadlec, P., Gabrys, B. & Strandt, S. Data-driven Soft Sensors in the process industry. *Comput. Chem. Eng.* **33**, 795–814 (2009).
10. K, O. *et al.* A Variable Scheduling Maintenance Culture Platform for Mammalian Cells. *SLAS Technol* **26**, 209–217 (2021).
11. Jones, D. R., Schonlau, M. & Welch, W. J. Efficient Global Optimization of Expensive Black-Box Functions. *J. Glob. Optim.* **13**, 455–492 (1998).
12. Qin, S. J. & Badgwell, T. A. A survey of industrial model predictive control technology. *Control Eng. Pract.* **11**, 733–764 (2003).
13. Sutton & Barto Book: Reinforcement Learning: An Introduction. http://incompleteideas.net/book/the-book-2nd.html.
14. Treloar, N. J., Fedorec, A. J. H., Ingalls, B. & Barnes, C. P. Deep reinforcement learning for the control of microbial co-cultures in bioreactors. *PLoS Comput. Biol.* **16**, e1007783 (2020).
15. Spielberg, S. P. K., Gopaluni, R. B. & Loewen, P. D. Deep reinforcement learning approaches for process control. in *2017 6th International Symposium on Advanced Control of Industrial Processes (AdCONIP)* 201–206 (IEEE, 2017). doi:10.1109/adconip.2017.7983780.

Acknowledgement

I would like to express my deepest gratitude to my supervisor, Prof. Makoto Taiji, for his invaluable guidance, support, and mentorship throughout the course of this doctoral research.

I am also deeply indebted to Prof. Genki Kanda and Dr. Miki Hojo, the co-first authors of the works presented in Chapters 2 and 3. I sincerely thank them for their collaboration, insightful discussions, invaluable advice, and constant encouragement. The opportunities I shared were essential to the success of this interdisciplinary project.

My sincere thanks go to my colleagues at Epistra: Dr. Yosuke Ozawa, Takeshi Sakurada, Prof. Koichi Takahashi, Tatsuki Higashi, Kensuke Tozuka, Takeshi Tatsukawa, Boy Jin, and Koki Kuroiwa. I am grateful for our shared mission to advance the societal implementation of CL-SDS and for their stimulating daily discussions. I also appreciate their consideration and support regarding my workload during the writing of this thesis. I also wish to thank Prof. Eiryo Kawakami and Takanori Kawabata at Chiba University for their understanding and accommodation regarding my research duties during this intensive writing period.

I own a special debt of gratitude to Prof. Osahiko Tsuji. Following the diagnosis of my spinal cord tumor last year, his rapid and skillful surgical intervention saved me from potential crisis. It was thanks to his care that I was able to recover and dedicate myself to writing this thesis.

Last but not least, I dedicate my deepest thanks to my family: my wife Yurisa, my daughter Eina, my father Shinji, and my mother Motoko. As the submission deadline of January 6th approached, they supported me through the New Year holidays, allowing me the time and space to concentrate solely on my research. Their sacrifices and enduring love are the cornerstone of my achievements.

I extend my appreciation to the researchers who provided critical feedback and technical support for the experiments detailed in Chapters 2 and 3. I thank T. Iwata, T. Mitsuyama, H. Ozaki, and N. Yachie for their insightful comments; H. Hirabayashi and E. Takagi for research support; A. Kato and H. Uchida for illustrations; and J. Freeman for carefully proofreading the manuscript. I am also grateful to all laboratory members at RIKEN BDR—specifically N. Koide, Y. Shibata, A. Maeda, and T. Maeda—for their kind help in the preparation of materials, their support in experiments, and their insightful discussions. Crucially, I would like to thank Mingming Zhao for providing the MYF5-tdTomato reporter hiPSCs and, along with Masae Sato, for valuable discussions on differentiation induction. I also acknowledge Aki Matsumoto and Megumi Goto for technical assistance, Kanae Mitsunaga for her constant help with FCM analysis, and Ayayo Sugiyama for assistance with data analysis.

研究実績

氏　　名（ふりがな）	都築　拓　　（　　つづき　たく　　）		
生年月日	1993年 3月 9日	現住所	非公開

学術論文名、著書、学協会、討論会、国際会議等での研究発表題名及び特許・発明等の名称	発表論文誌の名称、巻、頁、発表年（印刷中または投稿中の場合は、その旨を記入してください。） 発表学協会、討論会、国際会議等の名称及び発表年月日	共著者又は共同発表者氏名（本人を含め上から発表順に記入してください。）
Robotic search for optimal cell culture in regenerative medicine	Elife 11, e77007, 2022	Genki N Kanda, Taku Tsuzuki, Motoki Terada, Noriko Sakai, Naohiro Motozawa, Tomohiro Masuda, Mitsuhiro Nishida, Chihaya T Watanabe, Tatsuki Higashi, Shuhei A Horiguchi, Taku Kudo, Motohisa Kamei, Genshiro A Sunagawa, Kenji Matsukuma, Takeshi Sakurada, Yosuke Ozawa, Masayo Takahashi, Koichi Takahashi, Tohru Natsume
Automated evaluation of retinal pigment epithelium disease area in eyes with age-related macular degeneration	Scientific Reports 12 (1), 892, 2022	Naohiro Motozawa, Takuya Miura, Koji Ochiai, Midori Yamamoto, Takaaki Horinouchi, Taku Tsuzuki, Genki N Kanda, Yosuke Ozawa, Akitaka Tsujikawa, Koichi Takahashi, Masayo Takahashi, Yasuo Kurimoto, Tadao Maeda, Michiko Mandai
Different cell imaging methods did not significantly improve immune cell image classification performance	Plos one 17 (1), e0262397, 2022	Taisaku Ogawa, Koji Ochiai, Tomoharu Iwata, Tomokatsu Ikawa, Taku Tsuzuki, Katsuyuki Shiroguchi, Koichi Takahashi
Early and non-destructive prediction of the differentiation efficiency of human induced pluripotent stem cells using imaging and machine learning	Scientific Reports 15 (1), 26821, 2025	Miki Arai Hojo, Taku Tsuzuki, Yosuke Ozawa, Toshiyuki Araki, Hidetoshi Sakurai
Black-box optimization in immunology and beyond: A practical guide to algorithms and future directions	Allergology International 74 (4), 549-562, 2025	Takanori Kawabata, Taku Tsuzuki, Tsuyoshi Tatsukawa, Kota Matsui, Eiryo Kawakami
Production process optimization method and production process optimization system	US Patent App. 18/023,069, 2023	Y Ozawa, T Tsuzuki, T Higashi

Medium manufacturing method, medium manufacturing parameter determination method, medium, and program	US Patent 12,416,902, 2025	Y Ozawa, T Tsuzuki
Culture-related process optimization method and culture-related process optimization system	US Patent 12,291,754, 2025	Y Ozawa, T Tsuzuki, T Higashi
State inference system and state inference method	US Patent App. 18/571,492, 2024	Y Ozawa, T Tsuzuki
Linear Embedding-based High-dimensional Batch Bayesian Optimization without Reconstruction Mappings	arXiv preprint arXiv:2211.00947, 2022	SA Horiguchi, T Iwata, T Tsuzuki, Y Ozawa

（その他特記すべき事項）

上記のとおり相違ありません。

令和　8年　3月　25日

（氏　名）　都築　拓