

라벨 프리 비침습적 세포 이미징을 통한 미분화(undifferentiated) 마커 발현 수준 예측 기술

Ryutaro Akiyoshi*1

재생 의학 제품 생산 시, 미생물 오염 및 세포 손상을 방지하는 한편 세포 배양 과정과 품질을 지속적으로 모니터링할 필요가 있다. 이 연구에서는 기존의 파괴적인 방법을 사용하지 않고 비파괴적 세포 이미지 분석을 통해 줄기세포의 미분화 상태를 모니터링할 수 있는 새로운 방법을 개발하였다. 머신 러닝 알고리즘을 사용하여 물리적 접촉 없이 Oct3/4 및 Nanog와 같은 미분화 마커의 발현 수준을 예측하는데 사용할 수 있다. 이 방법은 명시적 현미경으로 획득한 iPSC 이미지와 정량적인 PCR 및 면역 염색법과 같은 파괴적 방법으로부터 얻은 평가 데이터에 기초한다. 비지도 및 준지도 모델을 적용하여 이미지 분석만으로 미분화 상태를 정확히 예측할 수 있다. 이러한 접근은 줄기세포 연구뿐 아니라 재생 의학 제품의 실시간 모니터링 및 새로운 품질 관리법에도 기여하고 있다.

개요

배아줄기세포 및 iPSC(induced pluripotent stem cells-유도만능 줄기세포)를 포함한 PSC (Pluripotent stem cells-만능줄기세포)는 세 가지 배아 발달 중배엽에서 유래한 모든 종류의 세포로 분화할 수 있다. 이러한 세포의 만능성 및 무한 증식 능력 덕분에 재생 의학, 약물 테스트, 질병 모델링 및 배아 발달 연구 등과 같은 다양한 분야에서 관심을 받고 있다. 재생 의학 제품의 원천으로서의 PSC는 미분화 상태를 유지해야 하고 미분화 상태로부터 벗어난 PSC를 검출하기 위해 마커 유전자와 단백질이 사용돼 왔다⁽¹⁾. 이러한 마커에는 미분화 세포에서 발현되는 세포 내 조절 인자(예를 들어, Oct3/4 및 Nanog)와 분화된 세포에서 발현이 감소되는 세포 표면 마커(예를 들어, SSEA-4 및 Tra-1-60)가 포함된다⁽²⁾⁻⁽⁴⁾.

그러나, 위의 분자 마커를 사용하여 세포의 미분화 상태를 평가하기 위해 FCM(flow cytometry-유세포 분석법), qPCR(quantitative polymerase chain reaction-정량적 중합 효소 연쇄반응), RNA-seq(RNA sequencing-RNA 시퀀싱) 및 면역 염색과 같은 파괴적인 방법이 필요하다. 재생 의학 제품의 제조 공정 중 비파괴적인 방법을 반복적으로 사용하는 경우, 제품에 사용돼야 할 세포의 누적 손실이 무시할 수 없을 정도로 커지게 된다. 이는 일반적으로 정기적인 장기간의 테스트가 요구되기 때문에 특히 세포 제품 제조 시 매우 중요한 문제이다⁽⁵⁾⁻⁽⁷⁾.

*1 Life Research & Development Department, Innovation Center, Marketing Headquarters

본 연구에서는 세포의 형태를 관찰하고 세포 상태의 변화를 탐지하기 위해, 세포 제품 제조 시 일반적으로 사용하는 명시야 현미경을 사용하였다. 명시야 현미경은 분석용 시약을 필요로 하지 않는 비파괴 방식이기 때문에, 세포 상태를 지속적으로 모니터링하는데 적합하며 비용 측면에서도 장점을 가지고 있다. 그러나 평가 결과는 관찰자의 경험과 인식에 따라 영향을 받기 때문에 일관성의 관점에서 문제점을 내포하고 있다. 이러한 문제를 극복하기 위해, 최근에는 사람의 판단에 의존하지 않는 명시야 현미경 이미지 분석에 대하여 연구하게 되었다. Christiansen 등은 iPSC 및 iPSC로부터 유래된 내피 세포에 발현된 CD31 단백질 형광 염색의 결과에 기초한 위상차 이미지로부터 CD31 단백질 발현을 예측하는 모델을 제안하였다⁽⁸⁾. 그러나 이러한 접근법은 형광 라벨 및 관련 전문가의 주석이 필요하다. Schmauch 연구진은 중앙 시료의 병리학적 이미지를 RNA-seq 데이터와 직접적으로 연관 지어 조직의 이미지로부터 암 관련 유전자의 발현을 예측하는 방법을 개발하였다⁽⁹⁾. 비록 이 방법은 세포의 형태를 식별하거나 개별 세포에 대하여 주석 처리할 필요 없이 높은 정확도를 구현할 수 있지만, 여기에는 셀을 염색하고 고정시키는 과정이 포함되기 때문에 여전히 파괴적인 방법이라 할 수 있다.

이에 우리는 세포를 염색하거나 전문가의 주석을 필요로 하지 않으면서, 살아 있는 세포에 대한 명시야 현미경의 결과에 기초하여 미분화 셀 마커의 발현을 예측하는 모델을 개발하였다. 우리는 비지도 및 준지도 모델을 통해 여러 분자 실험 결과를 정확히 예측하는 새로운 모델을 구축하고 검증하였다.

재료 및 방법

iPSC 준비

미분화 상태에서 편차의 원인이 될 수 있는 다음의 네 가지 조건을 사용하여 iPSC(201B7)를 배양하였다.

- (1)대조군(조건 1): iPSC의 표준 유지 배양 방법을 사용하였다.
- (2)저영양 조건(조건 2): 비활성화된 배양 배지(56°C에서 30분 동안 처리된 StemFit 배지)를 사용하였다.
- (3)분화 조건(조건 3): 10% FBS, 1% 최소 필수 배지 비필수 아미노산 및 1% GlutaMax가 포함된 Dulbecco's modified Eagle's medium을 사용하였다.
- (4)물리적 자극 조건(조건 4): 계대 배양 시 세포를 피펫으로 20회 서스펜딩한 후 iPSC를 위한 표준 유지 배양법을 사용하여 배양하였다.

각 조건에서 배양된 세포의 현미경 이미지 영상은 1일차부터 2일차까지, 그리고 4일차부터 5일차까지 촬영하였으며, 세포의 품질은 파괴적 방법(qPCR, FCM, 면역 염색, RNA-seq; 그림 1 참조)을 사용하여 평가하였다.

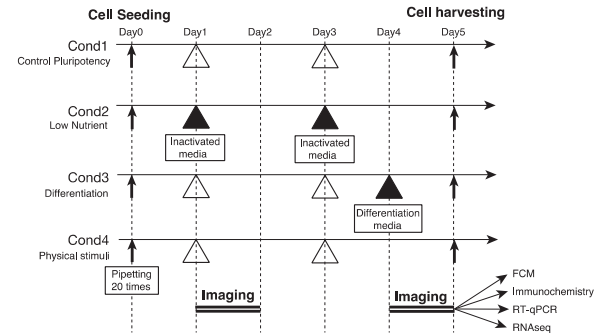


그림1 네가지 iPSC 배양 조건과 세포 평가 시점의 차이점
 흰색 삼각형: 공통 배양 배지 변경
 검은색 삼각형: 조건별 배양 배지 변경

iPSC의 명시야 이미지

6웰 플레이트에서 배양된 세포의 이미지는 20배 고배율 대물 렌즈를 사용하는 CellVoyager CQ1 (Yokogawa Electric Corporation)을 사용하여 획득하였다. 각 플레이트의 6웰 중 4개에서 5개의 영역이 5 × 8 사각형으로 구성되어 웰 당 총 200개의 시야를 제공한다. 각 필드에서, z축을 따라 5 μm 간격으로 5개의 평면 이미지를 획득하였다(그림 2 참조). 또한 총 20시간 동안 1시간 간격으로 타임랩스(저속 촬영) 이미지 작업을 수행하였다. CE(Contrast-enhanced-명암이 개선된) 명시야 이미지는 High Content Analysis Software System CellVoyager CellPathfinder(Yokogawa Electric Corporation)를 사용하여 다른 z축 위치에서 캡처한 이미지를 기반으로 생성되었다.

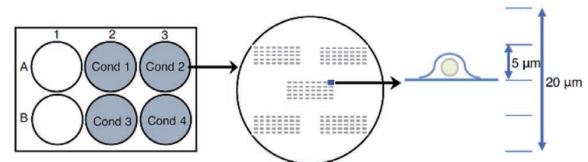


그림2 세포 이미징을 위한 6웰 플레이트 내 샘플 배치 및 각 필드에서의 데이터 획득 방법

파괴적인 방법의 결과를 예측하기 위한 모델의 생성

CE 명시야 이미지는 크기를 조정하고, 각 픽셀의 밝기를 조정하였다. 이러한 전처리된 이미지는 비지도 및 준지도 모델을 구축하는데 사용되었다. 준지도 모델은 One-Class SVM(one-class support vector machine)을 사용하여 구축하였으며, 파괴적인 기술(qPCR, FCM, 면역 염색, 및 RNA-seq)의 결과와 상관관계 계수 및 상관관계에 대한 평균 제공근 오차에 기반하여 서로 다른 특성을 가진 두 가지 모델을 도출하였다. Waisman 등의 분류 모델을 사용하여 네 가지 조건 하에서 배양된 iPSC 이미지를 분화된 상태와 미분화된 상태로 라벨링하는 가이드 모델로 활용하였다⁽¹⁰⁾. 분화 및 미분화된 iPSC를 예측하는데 높은 신뢰성을 가진 이미지를 추출하여 모델의 재학습 및 최적화를 위해 사용되었으며, 이를 기초로 하여 하나의 모델이 구축되었다.

결과

iPSC 이미지

5일차에 촬영된 iPSC의 명시야 이미지를 육안 검사하였을 때 조건 1, 2 및 4 중 차이점을 찾아내는 것은 매우 어려웠다(그림 3 참조). 그와 반대로 조건 3에서는 iPSC의 형태가 분명하게 비정상적(평평한 형태)으로 나타났으며 다른 조건에 비해 세포 밀도(confluency)가 현저하게 높았다.

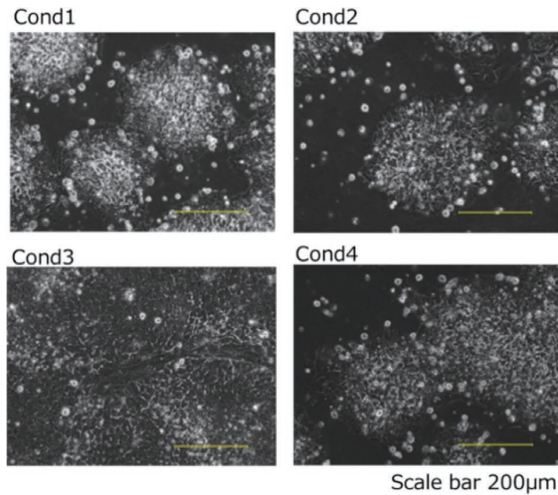


그림3 각 조건에서의 iPSC 명시야 이미지(5일차)
조건 1: 대조군; 조건 2: 저영양 조건;
조건 3: 분화 조건; 조건 4: 물리적 자극 조건

다양한 생물학적 데이터 분석

단백질 발현 분석(FCM)

FCM을 사용하여 미분화 세포에 대한 마커인 SSEA-4 및 Tra-1-602의 발현을 조사하였다. 미분화 세포의 비율은 전체 살아 있는 세포 중 수동 게이팅 또는 자동 게이팅(mindensity2)을 통하여 결정된 이중 양성 세포의 비율로 정의되었다. 미분화 세포의 평균 비율은 90%를 초과하였다(그림 4). 그림 4는 각 조건 하에서 미분화된 세포의 비율이다. 조건에 따라 미분화된 세포 비율의 차이를 시각화하기 위해 세로 축은 미분화된 세포를, 가로 축은 조건을 표시하였다. 그룹 간의 분산이 동일하지 않을 가능성을 고려하여 통계학적 유의성($p < 0.1$)을 판단하기 위해 Welch의 t 테스트를 사용하였다.

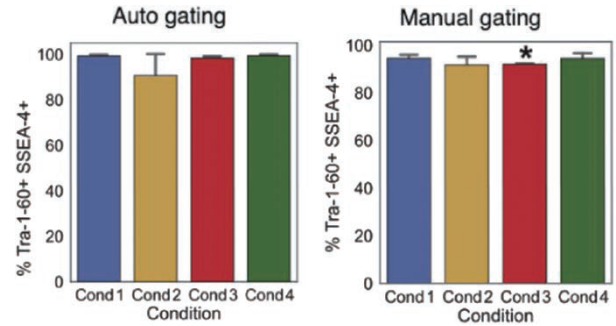


그림4 각 조건에서 미분화 세포의 비율은 FCM(자동 또는 수동 게이팅 사용)을 통해 측정됨
*: $p < 0.1$ (Welch의 t-검정)

단백질 발현 분석(면역 염색)

세포는 Oct3/4 및 Nanog에 대한 항체를 사용하여 염색되었으며, 양성 및 음성 염색은 BASC 알고리즘을 이용한 자동 임계값 설정과 수동 임계값을 설정하여 판단하였다. 각 임계값을 초과하는 강도의 양성 세포의 수를 기초로 미분화 세포의 비율을 계산하였다. 조건 1에서 Nanog 양성 세포의 비율은 수동 임계값을 사용할 때보다 자동 임계값을 사용할 때 더 낮았다. 자동 임계값을 사용한 경우 Nanog 양성 비율은 조건 1에서보다 조건 2 및 조건 3에서 유의하게 낮았으며($p < 0.05$), 특히 조건 3에서 크게 감소하였다. 반면 수동 임계값을 사용하는 경우, Nanog 양성 비율은 조건 1과 비교하였을 때 조건 2에서만 유의미하게 낮았으며($p < 0.1$), 이 프로파일은 수동 게이팅을 사용한 FCM으로 결정된 미분화 세포 비율과 유사하였다(피어슨 상관 계수, 0.93). 조건 1에서 Oct3/4 양성 세포의 비율은 자동 임계값을 사용할 때 수동 임계값을 사용할 때보다 더 낮았다. 자동화된 임계값을 사용하는 경우 Oct3/4 양성 세포 비율은 일반적으로 Nanog 양성 세포 비율보다 높게 나타났다. 자동 임계값 설정을 사용하는 경우 Oct3/4 양성 세포 비율은 조건 1에 비해 조건 2 및 조건 3에서 약간 낮았으며, 조건 1 및 2 사이의 차이는 유의미하였다($p < 0.1$). 반면, 수동 임계값 설정을 사용하는 경우 Oct3/4 양성 세포의 비율은 조건 2에서 조건 1보다 유의미하게 낮았고, 조건 3에서는 조건 1보다 유의미하게 높았다.

유전자 발현 분석(qPCR, RNA-seq)

Nanog 및 Pou5f1(Oct3/4 유전자)의 발현은 qPCR(그림 6A) 및 RNA-Seq(그림 6B)을 사용하여 분석하였다. 두 방법 모두에서 Nanog 및 Pou5f1의 발현 수준은 조건 2에서 조건 1보다 유의미하게 낮았다(그림 6A 및 6B 참조).

Nanog 발현 수준의 차이는 조건 1 및 2 사이에서 보다 조건 1 및 3 사이에서 더 크게 나타났으나 Pou5f1 발현 수준은 조건 1 및 3 사이에서 유의미한 차이가 관찰되지 않았다.

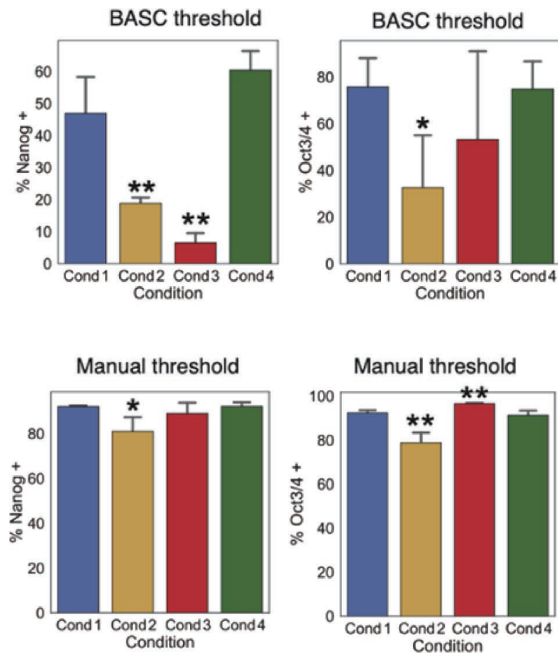


그림5 각 조건에서 면역염색을 통해 측정된 미분화 세포의 비율
위: BASC 알고리즘 기반 자동 임계값 설정 결과
아래: 수동 임계값 설정 결과
*: $p < 0.1$, **: $p < 0.05$ (Welch의 t-검정)

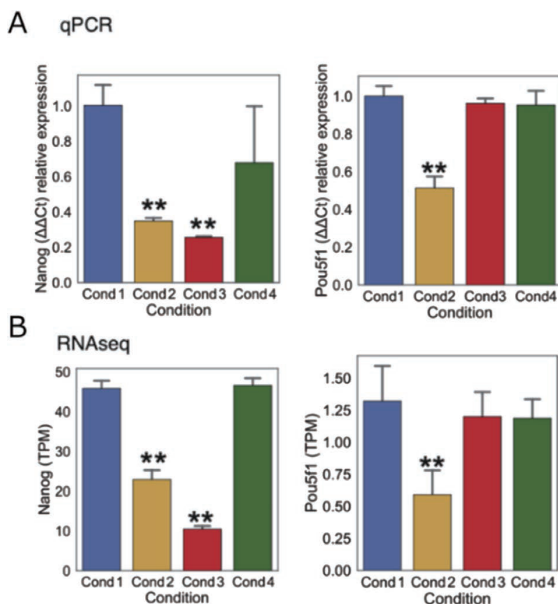


그림6 Nanog와 Pou5f1의 유전자 발현 수준 비교
A: qPCR, B: RNA-seq
*: $p < 0.1$, **: $p < 0.05$ (Welch의 t-검정)

모델에 의한 파괴 검사 결과의 예측

네 가지 배양 조건 하에서 구한 세포의 CE 명시야 이미지를 세 가지 모델에 입력하여 파괴 시험의 결과를 예측하였고, 미분화 세포 마커의 발현 수준을 예측하였다(그림 7 참조). One-Class SVM 프레임워크를 사용한 Model I은 Oct3/4에 대한 면역 염색 결과와 수동 임계값 설정(그림 5의 아래 패널, $r = 0.998$)에서 가장 강한 상관 관계를 보였고, 또한 Oct3/4 유전자에 대한 qPCR 및 RNA-seq 결과와도 강한 상관관계를 보였다. 모델 I 또한 자동 게이팅을 사용한 미분화 세포 마커 단백질의 FCM 결과와 강한 상관관계를 보였다(그림 4의 좌측 패널, $r = 0.912$). 동일한 One-Class SVM 프레임워크를 사용하는 모델 II는 자동 임계값 설정을 통한 Nanog 면역 염색 결과와 가장 강한 상관관계를 보였으며(그림 5의 위 좌측 패널, $r = 0.961$), 또한 Nanog 유전자 발현 수준과도 상관관계를 보였다(그림 6의 좌측 패널). 반면, 반지도 학습 프레임워크를 사용하는 모델 III는 자동 임계값을 사용한 Nanog 면역 염색 결과와 매우 강한 상관관계를 보였고(그림 5의 상단 좌측 패널, $r = 0.978$), Nanog 유전자 발현 수준과도 상관관계를 보였다(그림 6의 좌측 패널). 더구나, 모델 II와 모델 III는 수동 게이팅을 사용한 FCM 결과와 강한 상관관계(각각 $r = 0.908$ and $r = 0.938$)를 보여주었다(그림 4의 좌측 패널). 이 모든 상황을 고려해 볼 때, 각 모델은 미분화 세포 마커의 발현 수준을 높은 정확도로 예측하였다.

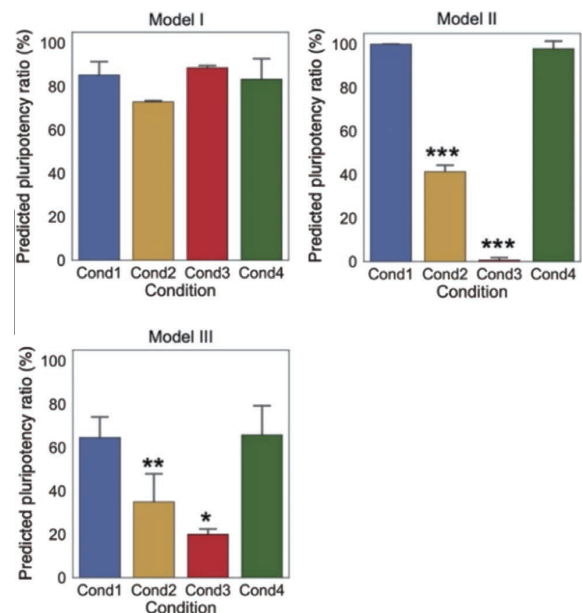


그림7 각 조건에서 파괴적 방법 결과를 예측하는
모델에 의해 예측된 미분화 세포 마커의 발현 수준
*: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$ (Welch의 t-검정)

논의 사항

후성유전학적, 전사적 및 대사적 상태와 같은 다양한 세포의 상태는 몇 시간, 몇 주 또는 그 이상의 긴 시간이 경과함에 따라 변화할 수 있기 때문에, 단일 마커나 단일 방법에 의존하여 검출하는 경우 이상 현상을 놓칠 수 있다⁽¹¹⁾⁽¹³⁾. 이 연구는 iPSC에서 미분화 세포 마커의 발현을 여러가지 파괴적 방법을 사용하여 자세하게 실험하였으며, Oct3/4 유전자 및 Oct3/4 단백질의 발현 사이에서 불일치를 발견하였다(그림 5 및 6 참조). 이는 세포 상태의 변화를 정확하게 포

착하기 위해 유전자 발현이나 단백질 발현에만 의존하는 방법의 한계를 시사하는 것이다. 본 연구에서 우리는 단백질 또는 유전자 발현을 각각 예측하는 독립적인 모델을 최적화하였고, 이와 같은 불일치가 있음에도 불구하고 세포 상태의 변화를 검출할 수 있다는 사실을 확인할 수 있었다. 다양한 기술을 사용하여 확인된 여러가지 분자 마커의 발현에서의 변화를 활용함으로써, 이러한 예측 모델들은 세포 상태를 종합적으로 포착하는 신호 감지 모델로서 최적화될 수 있다. 한걸음 더 나아가, 이 비파괴적 라벨 프리 접근 방식은 밀폐된 배양 환경에서 세포의 상태를 연속적으로 및 비파괴적으로 모니터링하는데 사용할 수 있고, 이는 이상이 발생하였을 때 손쉽게 공정 제어 또는 공정 중단에 대한 결정을 내릴 수 있게 만들어 줄 수 있다.

본 연구에서는 세포 이미지로부터 미분화 세포 마커의 발현을 예측하는 모델을 개발하였다. 이러한 접근법은 세포 집단 역학에 영향을 미치는 모든 분자의 발현에 있어 발생하는 변화(예를 들어, 세포 형태 및 세포 밀도)를 세포 이미지를 통해 예측할 수 있다는 사실을 보여주는 것이다. 이 모델을 세포 내 분자의 발현에 있어 변화를 검출할 수 있도록 확장함으로써 세포의 분화 상태에 대한 평가를 넘어 세포 연구 및 세포 제품 제조까지 가능하게 될 것이다. 또한 이 연구에서 개발된 모델 구축 워크플로를 홀로그래픽 이미징 및 RM-DIC(retardation modulated differential interference contrast - 지연 변조된 미분 간섭 대비) 현미경 관찰과 같은 새로운 이미징 기술과 통합하여 살아 있는 세포를 고해상도로 시각화하는 것이 가능해진다⁽¹⁴⁾⁽¹⁵⁾. 이렇게 함으로써 재생 의학 제품 제조에 있어 세포 품질 평가 작업을 위한 견고하고 신뢰성 있는 파이프라인의 구현 가능성이 확장될 것이다.

결론

우리는 iPSC 이미지로부터 비파괴적인 방식으로 높은 정확도의 세포의 유전자 및 단백질 발현 수준을 예측할 수 있는 모델을 개발하였다. 앞으로의 도전 과제에는 여러 세포주에 적용했을 때 모델 예측의 견고성을 입증하는 것과 분화된 세포에 대한 유사한 예측 모델을 구축하는 것이 포함된다. 우리는 이 기술을 실제 제조에 적용하기 위한 방법을 검증하고 우리의 고객 현장에서 그 가치를 시연하여 재생 의학 제품의 제조가 당면한 도전 과제에 대한 해결책을 제공하고자 한다.

참고문헌

- (1) P. W. Andrews, I. Barbaric, et al., "The consequences of recurrent genetic and epigenetic variants in human pluripotent stem cells," *Cell Stem Cell*, Vol. 29, No. 12, 2022, pp. 1624-1636
- (2) I. Chambers, J. Silva, et al., "Nanog safeguards pluripotency and mediates germline development," *Nature*, 450, 2007, pp. 1230-1234
- (3) J. A. Hackett, M. A. Surani, "Regulatory principles of pluripotency: From the ground state up," *Cell Stem Cell*, Vol. 15, No. 4, 2014, pp. 416-430
- (4) J. Silva, J. Nichols, et al., "Nanog is the gateway to the pluripotent ground state," *Cell*, Vol. 138, No. 4, 2009, pp. 722-737
- (5) A. Mamaeva, O. Krasnova, et al., "Quality control of human pluripotent stem cell colonies by computational image analysis using convolutional neural networks," *Int. J. Mol. Sci.*, Vol. 24, No. 1, 2023, p. 140
- (6) T. Kawase, K. Okuda, et al., "Non-invasive, quantitative assessment of the morphology of γ -irradiated human mesenchymal stem cells and periosteal cells using digital holographic microscopy," *Int. J. Radiat. Biol.*, Vol. 92, No. 12, 2016, pp. 796-805
- (7) M.-J. Kang, Y.-W. Cho, et al., "Progress in nano-biosensors for non-invasive monitoring of stem cell differentiation," *Biosensors*, Vol. 13, No. 5, 2023, p. 501
- (8) E. M. Christiansen, S. J. Yang, et al., "In silico labeling: Predicting fluorescent labels in unlabeled images," *Cell*, Vol. 173, No. 3, 2018, pp. 792-803.e19
- (9) B. Schmauch, A. Romagnoni, et al., "A deep learning model to predict RNA-Seq expression of tumours from whole slide images," *Nat. Commun.*, Vol. 11, 2020, p. 3877
- (10) A. Waisman, A. La Greca, et al., "Deep learning neural networks highly predict very early onset of pluripotent stem cell differentiation," *Stem Cell Rep.*, Vol. 12, No. 4, 2019, pp. 845-859
- (11) N. Rachinger, S. Fischer, et al., "Loss of gene information: Discrepancies between RNA sequencing, cDNA microarray, and qRT-PCR," *Int. J. Mol. Sci.*, Vol. 22, No. 17, 2021, p. 9349
- (12) E. Abruzzese, M. Bocchia, et al., "Minimal residual disease detection at RNA and leukemic stem cell (LSC) levels: Comparison of RT-qPCR, d-PCR and CD26+ stem cell measurements in chronic myeloid leukemia (CML) patients in deep molecular response (DMR)," *Cancers*, Vol. 15, No. 16, 2023, p. 4112
- (13) The International Stem Cell Initiative, "Assessment of established techniques to determine developmental and malignant potential of human pluripotent stem cells," *Nat. Commun.*, Vol. 9, 2018, p. 1925
- (14) I. Moon, B. Javidi, "Three-dimensional identification of stem cells by computational holographic imaging," *J. R. Soc. Interface*, Vol. 4, No. 13, 2007, pp. 305-313
- (15) K. Nishimura, H. Ishiwata, et al., "Live-cell imaging of subcellular structures for quantitative evaluation of pluripotent stem cells," *Sci. Rep.*, Vol. 9, 2019, p. 1777

* All company names, organization names, product names, service names, and logos that appear in this paper are either registered trademarks or trademarks of Yokogawa Electric Corporation or their respective holders.

라벨 프리 비침습적 세포 이미징을 통한 미분화(undifferentiated) 마커 발현 수준 예측 기술