

令和 4 年 6 月 21 日現在

機関番号：32684

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K07177

研究課題名（和文）細胞治療薬としての幹細胞の特性マーカー探索とAIによるバリデーション精度の向上

研究課題名（英文）Search for characteristic markers of stem cells for use in cell therapeutics and improvement of validation accuracy by AI

研究代表者

佐藤 光利 (SATO, Mitsutoshi)

明治薬科大学・薬学部・教授

研究者番号：60231346

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：足場非依存的コロニー形成の培養条件を検討し、0.03%ポリマーLA717を含む培地中に正常細胞(MRC-5, 1x10⁴ cells/well) および がん細胞 (HeLa-GFP, 1x10² cells/well) を低接着プレート上に共培養する条件が画像識別モデルに最適であることが示された。エッジ処理を行なった画像を用いた結果、バッチサイズ、学習率において最良の予測性能が認められ、エッジ処理の有効性が示された。本研究により、足場非依存的コロニー形成の細胞画像を人工知能(AI)を用いて解析することによって、簡便かつ効率的に異常細胞を検出可能になることが示唆された。

研究成果の学術的意義や社会的意義

本研究成果は、細胞加工製品による疾患治療への臨床応用と安全使用に向けて、幹細胞を生体に適用した際の細胞分化、再生能および組織適合性の細胞特性を明らかにすることに学術的意義がある。生体の組織修復に関わる幹細胞に着目した本研究は、目的細胞へ効率良く分化・増殖するための調節因子を遺伝学的および生化学的手法に加え、画像による細胞の形態学的特性を人工知能(AI)を用いて解析することで、高い精度の予測能から細胞バリデーションの精度を向上させ、実際に臨床応用する際の細胞加工製品の安全性を向上させるバリデーションシステムの構築、細胞加工製品の安全使用といった再生医療における社会的意義をもたらすものである。

研究成果の概要（英文）：To investigate the culture conditions for scaffold-independent colony formation, normal cells (MRC-5, 1x10⁴ cells/well) and cancer cells (HeLa-GFP, 1x10² cells/well) were co-cultured in a medium containing polymer LA717 at 0.03% on a low-adhesion plate to explore the optimization of a cell image identification model. Imaging using edge processing was able to recognize the optimal batch size and learning rate for performance prediction, and confirmed the effectiveness of the edge processing approach. Our findings suggest that abnormal cells can be detected easily and efficiently by analyzing cell images of scaffold-independent colonies using artificial intelligence (AI).

研究分野：生理学

キーワード：再生医療 医薬品安全性 幹細胞 細胞分化 バリデーション

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

間葉系幹細胞(Mesenchymal Stem Cell: MSC)は、生体の組織が障害を受けた際の組織修復に関わっていることから心筋再生療法における有力な治療ツールの一つである。間葉系幹細胞の虚血性心疾患治療に対する有効性を示すには、投与された細胞がおかれる虚血状態のストレス環境下での細胞の挙動を示すことが重要である。また、間葉系幹細胞を医療に用いる際は、細胞のソースとなるドナーやロットの違いによってもサイトカイン分泌能などの特性に差が生じることは治療効果にばらつきが生じることになる。したがって、細胞・組織加工製品の治療効果を治療開始前に保証するためには、組織修復能を予測するマーカーを見出すことが間葉系幹細胞を用いた際の治療効果を予測する上でも重要である。実際に、海外で行なわれた臨床研究では、十分量の心筋細胞が得られなかったことや適した表現型が得られなかったことによる有害事象発症の危険性などの問題が報告されている。

我々は、MSC がサイトカイン等の生理活性物質を分泌し、それらの生理活性物質が組織修復に関与していることを学会報告した(Satoh M. *et al.*, *J. Pharmacol. Sci.*, **115**, 172p, 2011)。また、心筋修復に関する研究では、マウス幹細胞を用いて細胞の分化・誘導メカニズム、心筋への選択的分化誘導法の開発を行い、心筋への分化を決定する遺伝子を DNA マイクロアレイによって発見し特許を取得した(特願 2006-109858 および特願 2012-209759)。さらに、異なる分化能を有する細胞株の特性解析と遺伝子発現プロファイリングデータを比較することにより細胞の分化能を予測するための候補遺伝子を見出した(Yasuda S. *et al.*, *Biochem. J.*, **437**, 345-355, 2011)。これらの研究によって、動脈硬化など血管病変の進展や制御、治療および予防対策ならびに血管や組織の修復機序の解明、幹細胞から心筋への選択的分化誘導法の確立に関して基礎となる成果を得てきた。

生体ではそれぞれの臓器への分化方向を決定するサイトカインが分泌されることから、MSC の分化・増殖をどのように *in vitro* で制御していくかが今後の臨床応用への課題である。しかし、MSC による心筋修復効果に関する実態は解明されておらず、しかも虚血性心疾患に対する有効性を示すには、投与された細胞がおかれる虚血状態のストレス環境下での細胞の挙動を解明することが重要である。また、MSC を細胞治療製品と考えた場合、ドナーやロットの違いによってサイトカイン分泌能に差が生じることは臨床問題であり、組織修復能を予測するマーカーを見出すことは MSC を用いた治療効果を保証する上でも重要である。したがって、これらに関係するパラクリン因子の探索や修復メカニズム、さらには、MSC から分化・増殖した心筋細胞・組織が、正常組織と変わらないことを保証することが臨床使用には重要になる。

2. 研究の目的

虚血性疾患である心筋梗塞の治療を目的とした医薬品としての MSC の臨床開発が進展してきていることから、細胞医療製品として MSC を使用する際には、製品の品質を評価するための細胞特性解析指標の設定、さらにそれらの指標による規格の設定がレギュラトリーの間からも重要になる。虚血性疾患治療に幹細胞由来細胞・組織を用いた試みがなされているが、分化・増殖した細胞を患者の体内に投与した際の有効性と安全性を保証するための明確な基準は示されていない。幹細胞由来細胞加工製品による虚血性疾患治療への臨床応用と安全使用に向けての基礎研究が本研究の目的であり、ヒト間葉系幹細胞の虚血条件下における遺伝子発現および生理活性物質濃度の変化を解析することにより生体適用時の間葉系幹細胞の分化・再生能および組織適合性等の細胞特性を明らかにする。本研究は、生体の組織修復に関わる幹細胞に着目して *in vitro* において間葉系幹細胞から目的細胞へ効率良く分化・増殖するための調節因子を遺伝学的および生化学的手法に加え、人工知能(AI)を用いて探索すること、また、細胞治

療薬を生体に適用した際に分化・増殖した細胞が正常細胞と変わらないこと、そして、組織適合性を有していることを生体への適用前に予測可能にすることにより、細胞加工製品としての有効性と安全性を保證するための遺伝学的・生化学的指標、ならびに AI による精度の高いバリデーションの確立を目的としている。

3. 研究の方法

(1) 虚血条件下における細胞培養

骨髄由来 MSC は、Lonza 社より入手した。虚血条件下における MSC の血管新生関連遺伝子発現変化を検討する目的で、MSC (継代数 9 (PS#9)) を 1×10^4 cells/well となるよう 96 穴細胞培養プレート上に撒き、MSCBM [Lonza] に MSCGM Single Quots [Lonza] を加えて調整した基本培地中で 37°C 、 $5\% \text{CO}_2$ の条件下で 80% コンフルエントになるまで培養した。その後、培養した上記の 96 穴細胞培養プレートを、コントロール群および虚血群に分け、次に示す通常条件または擬似的虚血条件においてさらに 24 時間培養した。

(2) 異常細胞混入モデル試験における細胞培養条件

ヒト胎児肺由来線維芽細胞(MRC-5)は、住商ファーマから購入し、GFP を導入したヒト子宮頸癌細胞株(HeLa-GFP)は、Cell Biolabs 社から入手した。これらの細胞は、10% ウシ胎児血清 (FBS; Sigma)、100 U/mL ペニシリンおよび 100 $\mu\text{g}/\text{mL}$ ストレプトマイシン(Nacalai Tesque, Kyoto, Japan)を添加した Eagle's minimum essential medium (MEM; Sigma, St. Louis, MO, USA) で維持した。細胞は、 37°C 加湿条件下、 $5\% \text{CO}_2$ および $95\% \text{N}_2$ で培養し、80% コンフルエンスに達したときに継代した。

正常細胞には MRC-5 (ATCC) を使い、がん細胞には HeLa-GFP (Cell Biolabs 社) を用いた。培地は、DMEM + 10%FBS + 1%NEAA + 0.03%LA717 (SphereMax) + 1%P/S (ペニシリン-ストレプトマイシン混合溶液) を使い、細胞を Corning 社低接着 96 ウェルプレート (#3474) に 1 ウェルあたり MRC-5(10,000 個) + HeLa-GFP(100 個)播種した (図 1)。

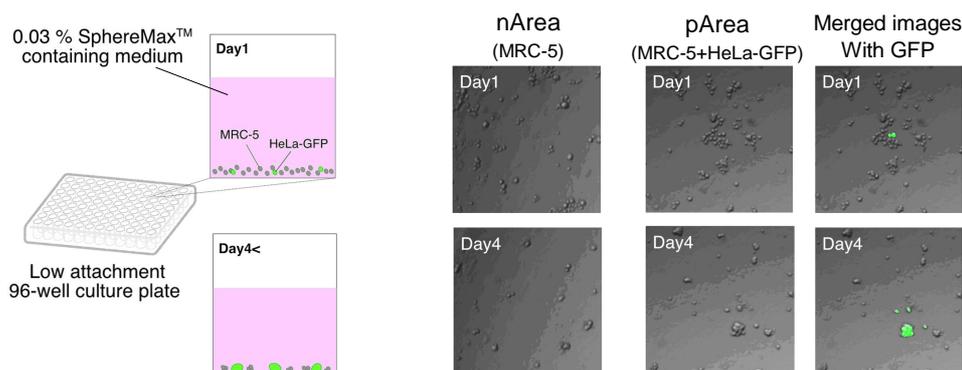


図1 細胞培養条件および解析画像に用いた画像例

(3) 画像解析に必要な細胞画像の撮像および画像解析方法

MRC-5 のみおよび MRC-5 + HeLa-GFP の共培養を 3 ウェルずつで実施し、各ウェルから 20 分おきに 4 日間 (96 時間) 連続的に画像を取得し (288 枚)、MRC-5 のみウェルから計 60 箇所 (20 箇所 \times 3 ウェル) を切り出し (negative area : nArea)、MRC-5 + HeLa-GFP の共培養のウェルから HeLa-GFP を含む領域 (positive area : pArea) を計 30 箇所 (10 箇所 \times 3 ウェル) を切り出して解析画像とした。

画像データについて nArea および pArea それぞれに 1 : 1 の割合で 2 群に分類し、Digits 入力用の Training および Validation データセットとしたものを "Total" データセットとした。また、"Total" データセットを細胞培養時間の時系列に基づいて培養経過時間の前半および後半

部分に 1 : 1 の割合で 2 群に分類し、前半部分を " Early_stage " データセットおよび後半部分を " Late_stage " データセットとした。

(4) AI を利用した培養細胞の異常細胞検出方法

がん化細胞のような目的細胞とは性質が異なった細胞を迅速に見つけ出すことは、細胞・組織加工製品を使用する再生医療では、重要になる。本研究では、AI を利用した培養細胞の異常細胞検出法の開発を試みることにした。最初の準備段階として、これまでの間葉系幹細胞からこれまでに分化誘導した心筋細胞の正常細胞の画像を取り込み、機械学習により正常細胞の解析を行う。異常細胞を人工的に添加して増殖させた培養組織を解析し、正常細胞の違いを AI に認識させることで識別因子を探索する。多くの細胞組織について繰り返し実施することにより、最終的には、ヒトによる確認では不可能なより精度の高い異常細胞の検出方法として確立させることを目的とする。

4 . 研究成果

(1) パラメータ調整

1 . バッチサイズ: ディープラーニングのように内部で確率的勾配降下法を使う場合学習する際に異常値の影響を小さくするために、データセットを幾つかのサブセットに分けて学習させた。この幾つかに分けたそれぞれのサブセットに含まれるデータの数をバッチサイズと呼び、本研究では、バッチサイズと細胞識別能力について検討を行なった。このとき、Training データセットの中から n 個のサンプルごとに分けて各集団に対して全ての勾配を計算し平均を算出する。その勾配によりウェイトを更新した後次の集団を利用して計算し、最終的に全ての集団を利用し終わった時点で Training データセットの順番をランダムに入れ替え、最初の集団に戻り次の学習を行った。

1-1. "Early_stage" データセット: 100 ~ 200 の 11 種類の Batch size について予測モデルを作製した結果、Batch size : 180 で最高性能 loss(val):0.986, Acc(val):40.0% となった。

1-2. "Late_stage" データセット: 5 ~ 200 の 24 種類の Batch size について予測モデルを作製した結果、Batch size : 189 で最高性能 loss(val):0.623, Acc(val):70.2% となった。

1-3. "Total" データセット: 5 ~ 350 の 26 種類の Batch size について予測モデルを作製した結果、Batch size : 30 で最高性能 loss(val):0.649, Acc(val):64.3% となった。

2. 学習率: 機械学習の最適化においてどのくらい値を動かすかというパラメータで、重みのパラメータを学習していく中で更新時の更新の幅を示すものである。ニューラルネットワークの学習過程において、入力データ (ベクトル x)、出力データ (ベクトル $y(x)$)、教師データ (ベクトル a) とした場合に、最終的に出力データと教師ベクトルとを等しくするために、ニューロンの重みと閾値を変動させて、下記の誤差関数 C をゼロに近似した。

$$\text{誤差関数 } C = 1/2 \| a - y(x) \|^2$$

2-1. "Early_stage" データセット: 0.01 ~ 10E-9 の 7 種類の学習率について予測モデルを作製した結果、学習率 : 10E-9 で最高性能 loss(val):0.682, Acc(val):65.9% となった。

2-2. "Late_stage" データセット: 0.01 ~ 0.00001 の 22 種類の学習率について予測モデルを作製した結果、学習率 : 0.0026 で最高性能 loss(val):0.612, Acc(val):70.6% となった。

3. Solver type 関数の最小化 (最大化) を数値解析的に最適化するアルゴリズム: ニューラルネットワークの学習過程において、損失関数の値を最小化するためにパラメータの更新が行われる。このパラメータの更新方法として、5 種類 (NAG, AdaGrad, Adadelta, Adam, RMSprop) のアルゴリズムの検証を行なった。最適解を見出すアルゴリズムとして、確率的勾配降下法

が一般に用いられるが、最近他の改良型が構築され精度の改善が報告されている。さらに、深層学習での学習においては、学習率の減衰 (learning rate decay)が必要で、5 種類の advanced learning rate (Step down, Exponential decay, Inverse decay, Polynomial decay, Sigmoid decay) についても同様に細胞識別能力について検討した。

3-1.“Early_stage”データセット: 5 種類の Solver type について予測モデルを作製した結果、Solver type : Sigmoid decay で最高性能 loss(val):0.611, Acc(val):70.2%となった。

3-2.“Late_stage”データセット: 5 種類の Solver type (NAG, AdaGrad, Adadelta, Adam, RMSprop) について予測モデルを作製した結果、Solver type : Adadelta で高い性能が観察されたため、さらに、この Adadelta を用いて Data transformation (image, pixel, none) について検証を行った。その結果、Adadelta および Data transformation (none) の条件において、最高性能 loss(val):0.678, Acc(val):65.9%となった。

本研究において、画像枚数については、現時点の精度であれば解析に十分な枚数であるが、更に予測精度を上げるには、より多くの情報を含む画像が必要と考えられる。今回の解析で培養細胞画像を培養期間の後半のグループで予測性能の向上が観察された。識別しやすい細胞特性が培養後半に出現していると考えられ、positive および negative 画像中に明らかに細胞特性に差が観られるデータセットを取得し、それらを用いることにより、これまで以上の学習・予測精度の向上が期待される。多くの単純な形状の細胞で学習・解析する際の過学習を改善するためには、インプットデータの精査を向上させることが必要である。また、機械学習にとって、どのような画像セットが最もモデル構築に適しているかの検討が必要である。左右、上下、および上下左右に反転させた画像について正しく特徴量を抽出できるか、さらには画像の拡大縮小によりノイズ混入による汎化性能への影響について検討が必要である。陽性および陰性領域の明確に異なる画像の取得により学習・予測性能の向上が期待される。エッジ処理などを行うことで細胞輪郭などの形態的特徴量をターゲットとしてインプットデータを収集し解析することにより学習・予測精度の向上が期待される。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Matsuzaka Yasunari、Kusakawa Shinji、Uesawa Yoshihiro、Sato Yoji、Sato Mitsutoshi	4. 巻 11
2. 論文標題 Deep learning-based in vitro detection method for cellular impurities in human cell-processed therapeutic products	5. 発行年 2021年
3. 雑誌名 Applied Sciences	6. 最初と最後の頁 9755 ~ 9770
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/app11209755	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 佐藤光利
2. 発表標題 ヒト幹細胞由来細胞組織加工製品の安全性に関する検討
3. 学会等名 第21回応用薬理シンポジウム
4. 発表年 2019年

1. 発表者名 草川森士、松坂恭成、植沢芳広、佐藤陽治、佐藤光利
2. 発表標題 足場非依存的コロニー形成試験における異常細胞コロニーの深層学習による画像識別モデルの評価
3. 学会等名 日本薬学会第141年会
4. 発表年 2021年

1. 発表者名 草川森士、松坂恭成、植沢芳広、佐藤陽治、佐藤光利
2. 発表標題 畳み込みニューラルネットワークを用いた異常細胞識別に関する検討
3. 学会等名 第22回応用薬理シンポジウム
4. 発表年 2021年

〔図書〕 計1件

1. 著者名 一般社団法人 日本医薬品安全性学会、宇野 勝次、佐藤光利 他	4. 発行年 2021年
2. 出版社 南山堂	5. 総ページ数 344
3. 書名 医薬品副作用・安全性ガイドブック	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	植沢 芳広 (UESAWA Yoshihiro) (90322528)	明治薬科大学・薬学部・教授 (32684)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------