

令和 3 年 6 月 4 日現在

機関番号：32612

研究種目：基盤研究(C)（一般）

研究期間：2018～2020

課題番号：18K06039

研究課題名（和文）CAGを超える恒常的プロモーターの新規同定

研究課題名（英文）Identification of novel and powerful constitutive promoter than CAG

研究代表者

中武 悠樹（NAKATAKE, Yuhki）

慶應義塾大学・医学部（信濃町）・講師

研究者番号：20415251

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）：生物の持つ分子メカニズムを解析するため、分子生物学の分野では、特定の遺伝子を誘導する実験が一般的である。遺伝子誘導には、プロモーターと呼ばれる配列を、目的遺伝子の上流に配置するが、体系的に有用な配列を探索する試みはされておらず、改善余地が大きいと考えられる。本研究では、最先端のオミックスデータをもとに、今まであまり着目されていなかった遺伝子の転写活性を検証した。本研究成果により同定されたプロモーター配列は、十分な活性を有していたが、古典的に使用されていたCAGプロモーターを超えなかった。これらのことから、配列そのものではなく、プロモーターの構造が重要であ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

古典的に使用されている分子生物学的ツールは、特段に不都合が無ければ、あまり着目されることもなく、盲目的に使用される傾向にある。ただ、実験目的によっては、既存のツールがうまく機能せず、得たい実験結果が得られない等のトラブルが生じる。本研究は、そのような申請者の実経験から着想され、最新の知見を駆使することで、古い固定概念を覆そうとした。今後、本知見を踏まえ、さらに良いツールの開発ができれば、今までできなかった実験評価が実施できるなど、生物学の分子メカニズムの解明への貢献が期待される。

研究成果の概要（英文）：To analyze molecular mechanisms, transgene induction is one of the most popular approach in biology field. Well-characterized promoters can induce the downstream gene of interest. Such promoters were established by previous studies, but systematic approaches have not been applied to identify the best candidates for it. In this study, transcriptome data was utilized for the identification of the candidate that could induce downstream genes. The identified gene promoters in this study retained a significant transcriptional activity, but not superior active than CAG promoter that is classically used in this field. These observations suggest that structure of promoter is more important than sequence themselves.

研究分野：分子生物学

キーワード：プロモーター 遺伝子発現 網羅的解析 恒常発現

様式 C-19、F-19-1、Z-19 (共通)

## 1. 研究開始当初の背景

本研究では、遺伝子を強制発現する際に必須な恒常的プロモータの新規開発をおこなう。現在、広く使われている恒常的プロモータは、ウイルス由来の CMV、SV40、ハウスキーピング遺伝子由来の EF1 $\alpha$ 、PGK、UBC、キメラ型の人工プロモータ CAG の 6 種が挙げられる。一番汎用性のあるとされるのは CAG で、これはヒトサイトメガロウイルス (CMV)、ニワトリ *Actin* プロモータ、ウサギ *Globin* イントロンを合成して作られた人工プロモータである。逆に、これら以外のプロモータは殆んど使用されず、恒常的プロモータに関する技術開発は 1980 年代で殆んどストップしており、網羅的発現解析が活発になった 1990 年代後半以降においても、体系的に恒常的プロモータを解析した報告は無い。

## 2. 研究の目的

下流遺伝子を恒常的に発現誘導するプロモータを新規に同定し、発現ベクターとして利用し、その有用性を分子生物学的に評価する。申請者は、独自に開発したデータベースを用いて、ヒト培養細胞等で高発現する遺伝子を網羅的に同定しており、そのうち 30 遺伝子のプロモータ活性を評価し、有用なプロモータを同定する。CAG などの分子生物学において汎用されているプロモータは、経験的に機能することが知られるだけであり、全転写産物を定量的に解析できる現代ならば、より有用なプロモータを同定できると考えられる。単なるバイオインフォマティクスの予想や分子ツールの開発に留まらず、網羅的データを起点としながら実証的なアプローチにより検証することで、真に実用的な情報を得る。

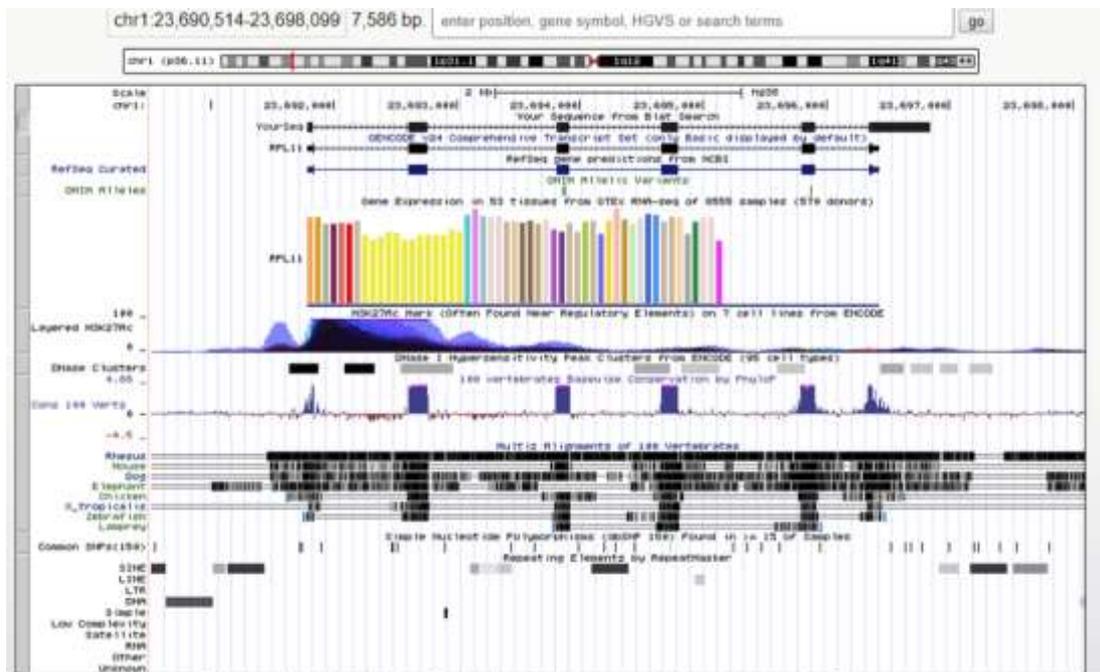
## 3. 研究の方法

平成 28 年度 4 月の時点で集積している別プロジェクトのオリジナルデータをデータマイニングし、30 遺伝子の候補を抽出する。それら遺伝子の転写開始点上流から約 2 kb をレポータープラスミド上に構築し、マウスおよびヒト ES 細胞にレポーターを導入する。ES 細胞を分化誘導することで、三胚葉での発現の安定性を培養条件下で検討する。高い転写活性を有するプロモータ(群)を同定したのちは、配列情報を In Silico で解析しつつ分子生物学的手法で実証することで、必要十分な遺伝子配列を絞り込む。これら分子生物学的な検討結果を反映させ、データマイニングの手法改善をおこなう。

## 4. 研究成果

### (1) 恒常発現している遺伝子の同定と候補遺伝子のプロモーター領域の解析

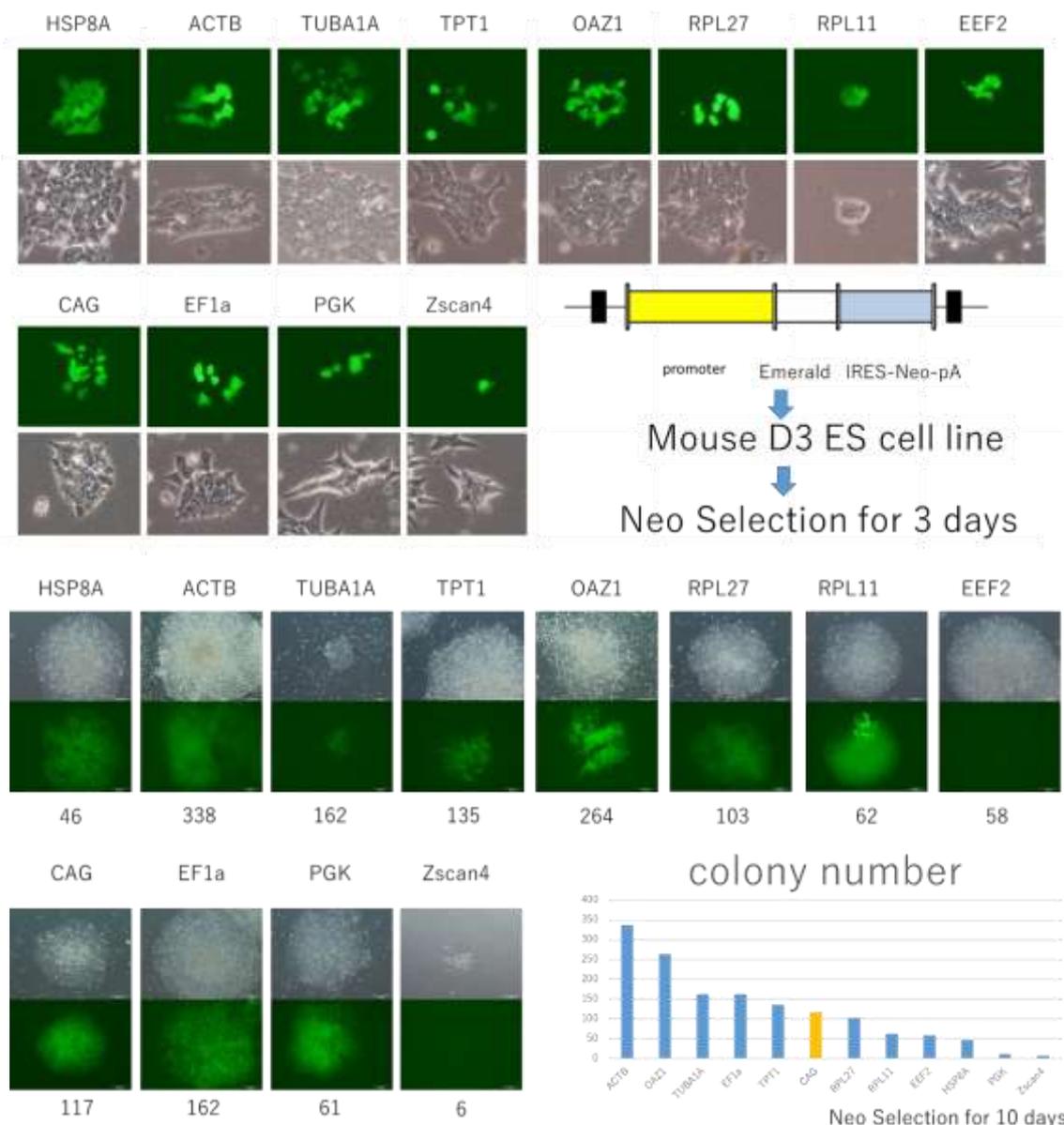
申請書に記載されている手順に従い、In-Silico にて、ヒト培養細胞等で高発現する遺伝子を網羅的に同定した。着手するに際し、Pseudogene の存在の有無が発現量に影響していることが懸念されたため、Pseudogene を 10 個程度有するような遺伝子は、候補から除外した。また、これらの恒常的に発現する遺伝子の転写開始点付近のクロマチン状態を解析したところ、ヒストン H3K27 のアセチル化が第一エクソンと第二エクソンのイントロン部分に集積していることが頻度高く観察された (8/10 遺伝子：下図は RPL11 遺伝子の例)。



H3K27 メチル化修飾は転写の活性化と密接に影響していると報告があり、イントロン構造を含む既製品プロモーター (EF1 $\alpha$ 、CAG) が高発現することから、第2エキソンに開始メチオニンが存在する遺伝子は、高アセチル化されたイントロンをふくみ、下流遺伝子の高発現への寄与が期待されるため、他の候補よりも優先順位が高いと考え、EEF2、TPT1、RPL27、RPL11、OAZ1、TUBA1BおよびHSPA8 遺伝子を、候補遺伝子群として選別した。

## (2) レポーター遺伝子の構築と活性の検討

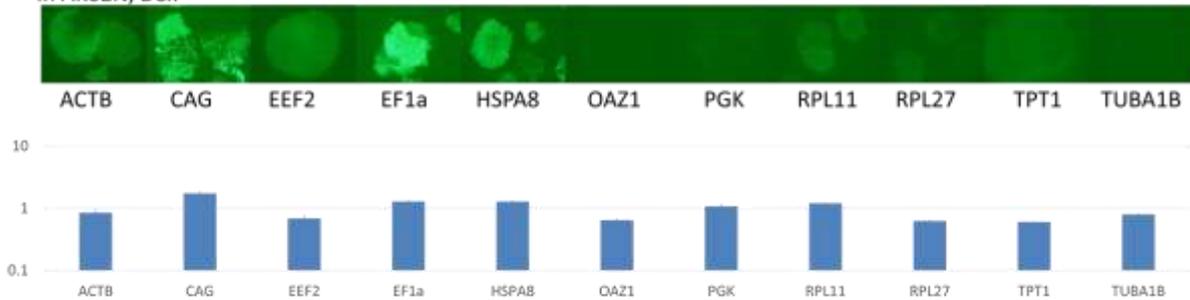
上述の候補遺伝子群について、転写開始点上流から 2 k b を取得した。これらのレポーター遺伝子コンストラクトを構築し、マウスおよびヒト ES 細胞に導入したところ、未分化な細胞状態では、CAG プロモーターと同等かやや劣る程度の転写活性を有していた。プラスミドの基本構造は、申請者の既報と同等である (下図および Nakatake et al. Cell Reports, 2020)。



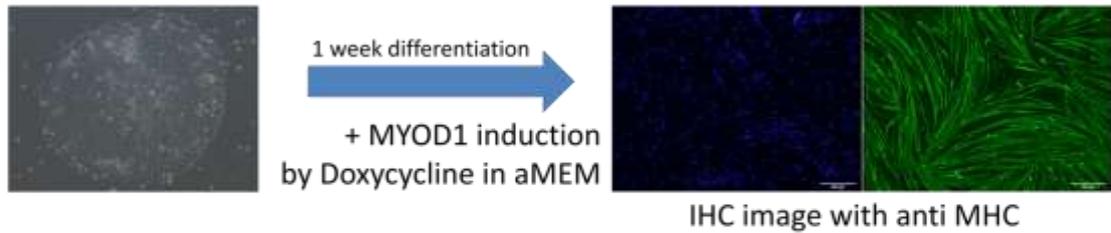
また、これらレポーターを成熟した筋細胞系に分化誘導できる MYO D1 誘導株に導入し、同様に転写活性を比較した。未分化状態と同様、これらの新規同定したプロモーターは転写活性を有していたが、比較対照となる CAG プロモーターより高い活性は得られなかった。一方、新規同定したプロモーターのうち、TPT1 および HSPA8 に関しては、CAG プロモーターよりも活性の変動が少なく、既存の分子生物学的ツールとして使用されている PGK プロモーターと比べると、活性が高いことがあきらかとなった (次ページ)。

IRES 配列を介して、薬剤選別をおこなっているため、上流の GFP 配列の mRNA は充分量の発現が得られているはずだが、蛍光が観察されていないレポーターがいくつか存在した (EEF2、TUBA1A、Zscan4)。QPCR で RNA の発現は確認できるため、翻訳に問題が発生しているものの、転写活性の強度の判定には影響がないため、蛍光での強度比較は未実施とした。

In AK02N, Dox-



Undifferentiated human ES cells



In aMEM, Dox-



In aMEM, Dox+



この結果、検討したすべての候補遺伝子のプロモーターは、CAG プロモーターよりも活性が同程度か弱いことが明らかになった。これは、NEUROG3 誘導後に神経細胞へと分化しても同様の傾向であった。一方、CAG プロモーターに比べ、TPT1 および HSPA8 のプロモーター活性は、NEUROG3 遺伝子誘導の有無で、活性の差が少なく、安定的なプロモーターであることが示唆された。

### (3) 考察と InSilico 解析手法の改善

研究計画では、CAG プロモーターよりも高レベルな発現強度を有するプロモーターを同定できると予想していたが、実際に検討した結果、安定的なプロモーターを取得することしかできていないため、いくつかの研究計画を変更する必要があった。CAG プロモーターは、プロモーター領域のみならず、CMV エンハンサー配列が持つ転写活性が上乘せされ、さらにウサギ Globin イントロンが連結されているが、このようなキメラなプロモーター構造そのものが、高発現するために重要であることが示唆された。

申請者の検討したレポーターも、イントロンを含有する構造を選別しているため、プロモーター領域に焦点をあてた実験系を構築するためには、1<sup>st</sup> Exon-Intron のつなぎ目付近の配列をスワップする対照実験系を組む必要があったが、期間内で実施することはできなかった。今後の課題としたい。

申請時点では得られていなかったが、申請者らが発表した論文報告 (Nakatake et al. Cell Reports, 2020) のデータを、最終年度では活用できたため、In Silico での解析を追加し、前年度までで解析した遺伝子以外の発現量に着目した。また、前年度までで得られた解析結果の再現性チェックをおこない、論文発表のためのデータを取得した。まず、In Silico の再解析については、申請時には得られていなかったデータセットをもちいて統合解析をおこなった。上述の申請者の発表論文より、ヒト E S 細胞において 510 の転写因子を誘導した 2135 のトランスクリプトームデータセットを抽出し、139 の転写因子を誘導したマウス E S 細胞のデータセット、申請時におこなったマウス GNF データセットおよび Illumina 社 BodyMap データセットを準備した。以前は、各データセットで発現強度の平均値が高いものをランク付けしていたのみであったが、更新した各データセットでは、中央値を重視してランク付けし、最大値および最小値の差をランク付けと加算する変更を施した。変更したデータテーブルの、上位 30 ランクの遺伝子には、ACTB および申請者が解析した 8 遺伝子を含むため、有意に改善されたデータテーブルと考えら

れた。今後、これらのデータを基に、前年度までに解析できなかった遺伝子について、同様の解析を進める必要があると考えられる。その際は、上述の Globin Intron を共通項目とした、プロモーター領域に焦点をあてた対照実験を構築すべきだろう。

#### (4) 結語

本研究は、CAG プロモーターを超える恒常的プロモーターを同定すべく、最新の網羅的な発現解析を基に、古き慣習を打破しようとした。しかし、逆に過去の偉人たちの業績を再認識する結果となった。申請者個人としては悔しい限りではあるが、最新のデータサイエンス的なアプローチの限界や、現行ツールの優秀さを、本研究が証明したともいえる。いずれ CAG を超える恒常性プロモーターが開発されると予想するが、本知見がその一助になることを願う。

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 0件）

|   |                               |
|---|-------------------------------|
| 1. 著者名<br>Nakatake Yuhki, (中略) Ito Noriko, Ishikawa-Hirayama Madoka, Mak Siu Shan, Jakt Lars Martin, Ueno Tomoo, Hiratsuka Ken, Matsushita Misako, Goparaju Sravan Kumar, Akiyama Tomohiko, Ishiguro Kei-ichiro, Oda Mayumi, Gouda Norio, Umezawa Akihiro, Akutsu Hidenori, Nishimura Kunihiro, Matoba Ryo, Ohara Osamu, Ko Minoru S.H. | 4. 巻<br>31                    |
| 2. 論文標題<br>Generation and Profiling of 2,135 Human ESC Lines for the Systematic Analyses of Cell States Perturbed by Inducing Single Transcription Factors  | 5. 発行年<br>2020年               |
| 3. 雑誌名<br>Cell Reports  | 6. 最初と最後の頁<br>107655 ~ 107655 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>10.1016/j.celrep.2020.107655  | 査読の有無<br>有                    |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難  | 国際共著<br>-                     |

〔学会発表〕 計1件（うち招待講演 0件/うち国際学会 0件）

|                                       |
|---------------------------------------|
| 1. 発表者名<br>中武悠樹                       |
| 2. 発表標題<br>CAGプロモータを超える新規恒常的プロモーターの同定 |
| 3. 学会等名<br>第41回分子生物学会年会               |
| 4. 発表年<br>2018年                       |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

| 氏名<br>(ローマ字氏名)<br>(研究者番号) | 所属研究機関・部局・職<br>(機関番号) | 備考 |
|---------------------------|-----------------------|----|
|---------------------------|-----------------------|----|

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 |
|---------|---------|
|---------|---------|