

平成 21 年 5 月 29 日現在

研究種目：若手研究 (A)

研究期間：2006 ~ 2008

課題番号：18681029

研究課題名 (和文) 種特異的選択的プロモーターの体系的解析による転写制御配列要素形成機構の解明

研究課題名 (英文) Characterization of Evolutional Turn-over of species-specific alternative promoters

研究代表者

鈴木 穰 (SUZUKI YUTAKA)

東京大学・大学院新領域創成科学研究科・准教授

研究者番号：40323646

研究成果の概要：motif expression decomposition (MED)モデルを用いて、プロモーター活性の予測可能性を検討したところ、相関係数で 0.8 程度の精度で予測を行うことが可能であった。10-fold のクロスバリデーションを行ったところ、本予測系は、データセットに対して、十分な再現性と許容範囲をもったものであることを確認することができた。また、プロモーター配列中にランダムな突然変異の導入系の開発を行い、約 5 % 程度の突然変異を簡便に導入することが可能となった。実際に EF1A1 を用いて、最大 10% の点突然変異の導入により、同プロモーターの活性が 0.03 倍から 6.5 倍まで改変できることを確認しており、用途に応じたプロモーターを創成することがある程度可能になったと考えている。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006 年度	7,900,000	2,370,000	10,270,000
2007 年度	7,300,000	2,190,000	9,490,000
2008 年度	7,300,000	2,190,000	9,490,000
年度			
年度			
総計	22,500,000	6,750,000	29,250,000

研究分野：複合新領域

科研費の分科・細目：ゲノム科学・基礎ゲノム科学

キーワード：プロモーター 分子進化 発現制御 完全長 cDNA 転写開始点

1. 研究開始当初の背景

多くの生物種は多様に变化する外的環境変化に対応するために、進化上、独自に新たな遺伝子発現制御システムを生成あるいは既存のシステムを改変してきた。相互にオルソログ関係にある遺伝子、またヒト・チンパンジー間といったごく近縁種の間であっても、その転写量が数倍程度異なる遺

伝子は何千も存在することを示唆する報告が相次いでなされている (Khaïtovich et al, Science, 2005)。また、結果として現存する種特異的転写制御システムの中には医学生物学上重要な発見あるいは産業上有用な応用につながる可能性をもつ特性を示すものも数多い。例えば、今年に入りドラフトゲノム配列の決定されたマカクザル CYP 遺伝子群にはヒトには存在しない多数の選択的プロモータ

ーが存在し、それら選択的プロモーターによる転写制御の多様化が、マカクザルにおいて様々な薬剤がヒトと異なる薬物代謝を示す遠因となって薬剤開発における前臨床試験の遂行上重大な障害となっている可能性が示唆されている。しかし、現在、10種類を超える哺乳類ゲノム配列が明らかになりつつある一方で、単純なゲノム配列比較からだけでは、いかにして進化上それらの転写制御様式が変化あるいは生成され、固定されたのかは、依然として明らかになっていない。

我々は、オリゴキャッピング法を用いた完全長 cDNA ライブラリーの作成法を開発し、完全長 cDNA 配列の収集と解析からゲノム上での転写開始点の正確な位置決定、さらにゲノム規模での上流プロモーター配列の同定と解析に取り組んできた (Suzuki and Sugano, *Methods Mol Biol*, 2003)。我々は、ヒトとマウスの代表的プロモーター配列の比較解析の結果、転写開始点近傍には配列保存性における不連続点が存在することを明らかにし、進化の過程で個々の塩基置換とともに塩基配列の大規模な変化が、転写制御に関わると考えられる領域内においても比較的頻繁に行われてきた可能性を示してきた (Suzuki et al, *Genome Res*, 2004)。さらに、これまでに蓄積してきた 180 万の完全長 cDNA 配列 5' 端塩基配列の解析から、ヒトゲノム中には代表的プロモーターに加えてこれまでに見出されていなかった選択的プロモーターが多数存在し、ヒト遺伝子全体の約半数が選択的プロモーターにより制御されていること明らかにした (Kimura et al, *Genome Res*, 2006)。これらの結果は最近になって提出された、転写制御機構が従来考えられていたよりも遥かに多様に進化しているという仮説を支持するものである。しかし、我々が新たに見出した多数の選択的プロモーターについてもマウスとの配列比較解析を行ったところ、多数の完全長 cDNA (転写開始点データ) に支持されるものに限っても、対応する配列自体がマウスゲノムに存在しないものが大部分であることが明らかとなった。またこれらのヒト特異的選択的プロモーター候補群は、進化上保存される選択的プロモーターと比べて、1) GC 含量が低く CpG アイランドに乏しい、2) TATA ボックス様配列に富む、3) 転写活性化能として属する遺伝子中で最も強力なプロモーターであることはほとんどない等の特徴を顕著に共有していることが明らかになっている (Tsuritani et al, *Genome Res* 2006)。

2. 研究の目的

本研究は、基本転写因子あるいはその他の転写調節因子と DNA との直接の相互作用の場となる転写開始点近傍配列ゲノム DNA 中 (転写開始点に対して上流 1kb、下流 200bp 程度を対象とする) に生じた大小の塩基置換のどの部分が、いかにしてプロモーター活性の獲得と変化に本質的な役割を果たしているのかを、その転写活性化能の種間比較により明らかにすることを目的とする。ヒト特異的プロモーターを主たる標的として、体系的なルシフェラーゼレポーターアッセイを行い、転写活性可能とプロモーター配列の相関についての解析を行うものである。

ヒト特異的選択的プロモーター候補の体系的な解析からゲノム配列中でのプロモーターの生成あるいは転写活性化能の改変メカニズムの解明に突破口をさぐる。また、ヒト特異的プロモーター候補に関して、現在ゲノム配列が入手可能となっているチンパンジー、マカクザル、ウシ、イヌ、ラットを用いて、分子進化系統樹を順次さかのぼり、どの段階で該当する配列が生成されたのかを同定、解析することにより、さらに、多面的な解析を行う。それぞれのゲノム領域を PCR によりクローニングしルシフェラーゼレポーターアッセイを行い、いかなる塩基置換がプロモーターとして転写活性可能を獲得するのかに本質的であるのかを明らかにする。

3. 研究の方法

転写活性化能がヒトプロモーター、対応するマウスゲノム間でどの程度異なるかについて定量的、体系的にルシフェラーゼアッセイを行った。約 100 個のヒト種特異的選択的プロモーター候補と対応するマウスゲノム配列に対して、プライマーを設計し、それぞれのゲノム DNA を PCR により増幅、クローニングを行った。クローニングは、汎用クローニングシステム Gateway を用いて 96 穴プレートにて行い、得られるクローンは両端からのシーケンス確認の後、将来の多目的的利用を見込んでグリセロールストックの遺伝子資源として保存することとした。単離された物理的クローンを用いて、ルシフェラーゼレポーターアッセイによる転写活性化能の測定を行い、転写活性化能がヒトプロモーター、対応するマウスゲノム間でどの程度異なるかについて定量的な解析を行った。活性測定には汎用的培養細胞系としてヒト 293 細胞、およびマウス NIH3T3 細胞を用い、細胞環境あるいは転写因子側の種間変化の影響を評価、極力排除するものとした。転写活性可能と塩基

置換位置の相関関係を検証することにより、転写活性化の獲得に本質的変異部位の同定を試みた。また、ゲノム中に塩基置換が導入された場合、そのプロモーター活性化能のランドスケープがどのように変化するかについて、現在公開されている 20 種類の哺乳類のゲノム配列を用いてプロモーター活性予測に関するシミュレーションを行った。

4. 研究成果

ヒトプロモーター領域約 450 種類、ランダムに抽出したゲノム領域 250 種類、マウスについてその対応する遺伝子 100 種類について収集したプロモーター活性データを用い、プロモーター活性とプロモーター配列の相関関係の検証を行った。motif expression decomposition (MED) モデルを用いて、プロモーター活性の予測可能性を検討したところ、相関係数で 0.8 程度の精度で予測を行うことが可能であった。また、10-fold のクロスバリデーションを行ったところ、本予測系は、データセットに対して、十分な再現性と許容範囲をもったものであることを確認することができた。

一方で、ヒト-マウス間での配列比較から、対応するゲノム配列は存在するものの近傍に転写開始点データがマウス側にのみ存在しないものが約 1000 種類同定した。その中には、遺伝子全体としては多くの転写開始点に対応しているものの、該当する選択的プロモーターに限って転写開始点が存在せず、ヒトにおける転写開始点データの選択的プロモーター間での振り分けを指標とした統計的推定から ($p < 0.001$)、マウスの転写開始点データの coverage 不足に原因が求められないと考えられるものが約 250 種類含まれていた。またこれらの中にはヒトでユビキタな発現パターンを示すものが約 100 種類見出された。これら 100 種類をクローニング、探索的に転写活性化能の測定を行い、有為に転写活性化能の差を示すもの 50 種類を見出した。これらについても、同様に MED による解析を行ったが、MED を用いた系ではそのプロモーター活性の相違を説明することができなかった。また、マウスモデルから構築した MED でヒトプロモーター活性を十分な精度で予測することも多くの場合、困難であった。

活性の進化的変遷を比較する一連の実験において、少なくとも一群の遺伝子については予想される以上の相関の低さの原因が不明であったことから、測定系において、用いているレポーター遺伝子の翻訳レベル

での発現調整の効率の影響が懸念された。そこで、これまでのベクター系を改変し、(Internal Ribosomal Entry Sequence) を用いたレポーター系に移行することで、より正確に一次配列の転写活性可能を測定できる系を構築した。保有するほぼ全てのクローンについて、ベクター系の改変が終了しているものの、研究期間中に全実験データの検証を行うには至らなかった。

我々は同時にプロモーター配列中にランダムな突然変異の導入系の開発を行った。配列改変によるプロモーター活性への影響を評価するためには、迅速簡便に配列中にランダムに塩基置換を導入する方法の開発が必須である。従来の PCR を改変した系を利用した実験系を用いることにより約 5 % 程度の突然変異を簡便に導入する方法を開発することが可能となった。

興味深いことにこの系を利用して、既知のプロモーター配列に対して、ランダムに突然変異を導入することで、天然のプロモーターと異なったプロモーター特性をもった配列を創出することが可能であった。現在、組み換えタンパク質産生系においては、いわゆるハウスキーピング遺伝子のプロモーターが用いられることが多い。しかしタンパク質によっては発現の程度によっては活性が喪失するもの、あるいは特定の細胞でタンパク質を発現させた方がよいもの等、現在、利用可能なプロモーター群では必ずしも十分に多様な用途に対応することはできない。本研究で確立された、プロモーター改変-検出系を体系的に用いることにより、用途に応じたプロモーターを創成することがある程度可能になったと考えている。実際、EF1A1 を用いた予備実験から、最大 10% の点突然変異の導入により、同プロモーターの活性が 0.03 倍から 6.5 倍まで改変できることを確認している。

また、ランダムな変異導入系から得られた人工的プロモーター群は、MED によるプロモーター活性予測についても新たな角度からの教育データを与えると考え、5 個の遺伝子についてランダムに突然変異を導入した 100 種類の人工プロモーターの活性測定と解析を行った。同様に MED モデルによる評価を行ったが、天然に存在するプロモーター活性予測とは MED におけるモデルの安定性、その他のパラメーター群が大きく異なっていた。現在、ヒト、マウスプロモーターとその進化的変化、あるいは人工的に合成したプロモーターの予測精度が大きくその性質を異にする理由について計算機的な解析を進めている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 2件)

Davuluri, R.V., Suzuki, Y., Sugano, S., Plass, C., Huang, T.H.

論文名: The functional consequences of alternative promoter use in mammalian genomes

掲載雑誌: Trends Genet 巻: 24

最初と最後の頁: 167-77 発行年: 2008

査読の有無: 有

Sakakibara Y, Irie T, Suzuki Y, Yamashita R, Wakaguri H, Kanai A, Chiba J, Takagi T, Mizushima-Sugano J, Hashimoto S, Nakai K, Sugano S

論文名: Intrinsic promoter activities of primary DNA sequences in the human genome.

掲載雑誌: DNA Research 巻: 14

最初と最後の頁: 71-77 発行年: 2007

査読の有無: 有

6. 研究組織

(1) 研究代表者

鈴木 穰 (SUZUKI YUTAKA)

東京大学・大学院新領域創成科学研究科・准教授

研究者番号: 40323646

(2) 研究分担者

(3) 連携研究者