

科学研究費補助金研究成果報告書

平成 24 年 10 月 19 日現在

機関番号：82610

研究種目：基盤研究 (B)

研究期間：平成 21 年度～平成 23 年度

課題番号：21390283

研究課題名 (和文) 膵β細胞の新規転写調節機構と機能性 RNA の解析

研究課題名 (英文) A study on regulatory genomic regions and functional RNAs in pancreatic β-cells.

研究代表者 安田和基 (KAZUKI YASUDA)

独立行政法人国立国際医療研究センター

研究所 糖尿病研究センター 代謝疾患研究部 部長

研究者番号：80311611

研究成果の概要 (和文) : 高度に機能分化したインスリン産生膵β細胞における遺伝子発現調節機構を解明するために、ラット膵β細胞株 INS-1D 細胞から、Vector-capping 法を用いて全長 cDNA ライブラリーを作成した。26,976 クローンについて 5'-側 one pass シークエンスを行い、データの品質チェック後 6,042 クラスターを得た。膵β細胞機能に関わるいくつかの遺伝子について、既報と大きく異なる転写開始点が同定され、転写因子結合予想配列や比較ゲノム学的見地から検討を行った。また明らかな ORF (Open reading frame) をもたず、インスリン分泌に関与しうる、機能的な non-coding RNA (ncRNA) も存在した。

研究成果の概要 (英文) : Insulin producing pancreatic β-cells are highly differentiated tissue, In order to elucidate molecular mechanisms of transcriptional regulation in β-cells, we constructed a novel full-length cDNA library by Vector-capping method from rat β-cell line INS-1D cells, and performed a 5'-one pass shotgun sequencing with 26,976 clones, resulting in 6,042 clusters. We identified novel transcriptional start sites (TSSs) and possible transcriptional regulatory regions for some genes. We also obtained several functional non-coding RNA (ncRNA) candidates, which may be involved in insulin secretion. Thus, our system should serve as a novel tool for investigating 'RNA world' of pancreatic β-cells.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
平成 21 年度	6,600,000	1,980,000	8,580,000
平成 22 年度	4,000,000	1,200,000	5,200,000
平成 23 年度	3,300,000	990,000	4,290,000
総計	13,900,000	4,170,000	18,070,000

研究分野：医歯薬学

科研費の分科・細目：内科系臨床医学・代謝学

キーワード：糖尿病、遺伝子、発現調節、RNA、膵β細胞

1. 研究開始当初の背景

- (1) ヒトゲノム配列解読後の、いわゆるポストゲノム研究において、ゲノムのほとんどの部分が転写されていることや、多彩な alternative splicing の存在など、従来の予想よりはるかに複

雑かつ多様な転写産物の世界が発見され、「RNA world」という呼称が登場しつつあった。中でも明らかなタンパクをコードしない、非コード性 RNA (non-coding RNA、以下 ncRNA) が注目を集めはじめていたが、特に高度に

機能分化した組織における、生理的意義や病態との関係は、ごく一部の例外を除き、ほとんど不明であった。

- (2) 研究代表者は、インスリンを分泌する膵β細胞の分子生物学的解析とともに、ヒト2型糖尿病の遺伝解析に従事してきたが、本研究開始直前に報告 (Yasuda K et al Nature Genetics 2008) した *KCNQ1* 遺伝子をはじめ、多くの遺伝因子が、イントロンや遺伝子間領域に存在しており、膵β細胞における「ゲノム生物学」のさらなる理解が必須と考えた。
- (3) 膵β細胞は、グルコース反応性インスリン分泌 (以下 GSIS) という機能特性を有するなど、きわめて高度に機能分化した臓器である。膵β細胞特異的な遺伝子やバリエーションも多く、その特徴的な遺伝子発現調節機構の解明は、代謝制御システムや、糖尿病の病態の解析においても、きわめて重要である。マイクロアレイによる研究は行われていたが、より網羅的な発現遺伝子解析は行われていなかった。

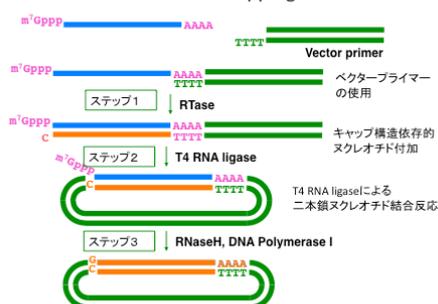
2. 研究の目的

本研究では、インスリン分泌臓器である膵β細胞について、我々が世界にさきがけて作製した完全長 cDNA ライブラリーを活用し、特異的な遺伝子発現調節機構、及び ncRNA を含む新規 RNA と細胞機能との関連を検討する。これにより高度に機能分化した細胞における、これまで知られていなかった「RNA world」の機能的な役割を明らかにし、さらにその *in vivo* における生理的・病的意義を検討する。

3. 研究の方法

- (1) 培養膵β細胞株 INS-1 細胞の完全長 cDNA ライブラリー由来クローンのシーケンス決定
 - ① 完全長 cDNA クローンの 5' -側 one pass シークエンス: 我々は、生理的な膵β細胞に性質に近いとされるラット INS-1D 細胞 (Dr Wollheim、及び関根信夫博士より供与) から、Vector-capping 法 (DNA Res 12:53-62, 2005) にて、完全長 cDNA ライブラリーを作成した (図 1)。

図 1: Vector capping 法



この cDNA ライブラリーを、electroporation 法にて大腸菌に導入、トランスフォーメーションののち、シングルクローンをランダムピックし、粗精製したプラスミドについて、ベクター (pGCAP-1 及び pGCAP-2) 由来のプライマーを用いて、ABI 社製 3730XL により、5' -側 one pass シークエンスを行う。

- ② クオリティーチェック (QC): シークエンスの波形を Phred で数値化した後、マスクング、ベクター部分のトリミング、ポリ A の存在確認を経て、信頼できるデータとして以降の解析へすすむ。
- ③ 完全長クローンか否かの判定 (ライブラリーの評価): 本法では、逆転写 (RT) 反応の際、5' 端 Cap 構造依存的に、ヌクレオチド (主に C) が付加される。この性質を用いてマッピングしたゲノムと比較すれば、完全長の可能性が高いかどうか、判定が可能である。
- ④ 配列のクラスタリング、及びアノテーション: クローン配列からクラスタリングを行った後、RefSeq データベースへ相同性検索 (BLAST) を行う。

(2) 膵β細胞の転写調節領域 (候補領域) の *in silico* 解析

完全長 cDNA の利点の 1 つは、転写開始点が判明することにより、ゲノム上の転写調節領域が同定できることである。NCBI のデータベースあるいは既報と比較し、より上流に転写開始点が同定されたものについて、その 5' 上流域 1-3kbp を「新たな転写調節領域候補」として、まず *in silico* で転写調節エレメントの予測を行う。哺乳類細胞のプロモータ配列に関するソフトウェアとしては ElDorado、GEMS Launcher など (独 Genomatix 社) を利用する。また、ヒト、マウスなどの種間のゲノムで転写開始点及び転写調節領域の保存性を検討する。

(3) 膵β細胞における ncRNA の機能および発現解析

500bp 以上のインサートをもちながら明らかな ORF を持たないクローンと「non-coding RNA (ncRNA)」候補として検討する。

- ① ノックダウンによる機能的変化: 配列特異的な siRNA を複数デザインし、INS-1D 細胞に導入して、当該 RNA の発現を抑制 (ノックダウン) する。次に、当該 RNA をノックダウンした INS-1D 細胞について、バッチインキュベーション法により、3mM および 25mM グルコースに対するインスリン分泌に変化が見られるかを検討する。
- ② 機能性 RNA の分子メカニズムの解析: 前記機能的な意義が想定されたクローンに

については、siRNA 処理後に INS-1D 細胞から RNA を回収し、Affymetrix 社 GeneChip システム (Rat 230 2.0 アレイ) を用いて、網羅的な遺伝子発現プロファイルへの影響を調べ、その分子メカニズムを検討する。

- ③ *in vivo* での発現解析：糖尿病モデルラット (GK ラット) 及び対照ラット (Wistar ラット) において、短い相補的な RNA プローブを用いて脾における ISH (*in situ* hybridization) を行い、発現パターンを解析する。解析プローブと相補的な RNA (得られたクローンと同じストランドで便宜的にセンス鎖 RNA と呼ぶ) を、ネガティブコントロールのプローブとする。

4. 研究成果

- (1) 培養脾 β 細胞由来全長 cDNA ライブラリーのクローン解析

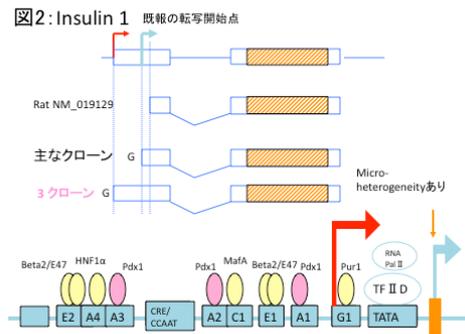
単一コロニーからピックアップした合計 26,976 クローンについて、5' 側の one pass シークエンスを行った。クオリティチェックの後、17,276 クローン (64.0%) を readable data として以降の解析へ進めた。パイロット的に抽出した約 1,000 クローンについて、「完全長 cDNA」と判定されるクローンは 87%、厳しい基準でも約 70% にのぼった。全クローンは 6,042 クラスターに分類され、うち 1 クローンからなるクラスターは、3,582 であった。

遺伝子アノテーションをみると、クローン数の多い遺伝子としては、複数のクラスターにわたったインスリン (Ins1: Insulin1 及び Ins2: Insulin2) が第 1 位で、他に Iapp (Islet amyloid polypeptide) など脾 β 細胞に特異性の高い遺伝子、Chga (Chromogranin A)、Chgb (Chromogranin B)、Ppia (peptidylprolyl isomerase A) など内分泌細胞に特徴的なタンパク、Gapdh やリボソームタンパクをコードする遺伝子群、ミトコンドリア機能に関わる遺伝子群、housekeeping 遺伝子が並び、予想通り発現量の多い遺伝子が上位にみられた。

- (2) 転写開始点 (TSS) の同定と、機能的な転写調節領域候補の抽出

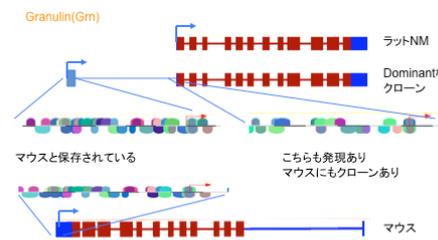
5' 側 one pass シークエンスから得られた配列についてデータベース検索を行った。

- ① Ins1 (Insulin1) (図 2)：インスリン 1 の TSS は、図に示すように既報の転写開始点で比較的良好に一致していたが、数塩基の規模での変動はあった。これは TSS のいわゆる「microheterogeneity」と呼ばれる良く知られた現象であり、生物学的意義は乏しいと思われる。



- ② Chgb (Chromogranin B)：この遺伝子では、当時 RefSeq の NM クローン (NM_012526) に比べ、数十 bp~110bp 上流からスタートすると思われる、複数のクローンが存在した。これらはいずれも 5' -UT が長くなるが、最も 5' 側の TSS 候補とマウスの既報転写開始点とが、非常に良く対応することから、これがラットでも TSS の「本命」である可能性が高い。
- ③ Iapp (Islet amyloid polypeptide)：既報 TSS とほぼ一致するクローンが多いほか、逆により下流からスタートするクローンがみられた。ただしクローン数は限られ、全長 cDNA に至らなかったクローンの可能性も高い。
- ④ このほかにも既知遺伝子だが既報より 500bp 以上上流に転写開始点 (TSS) が同定されたものも存在した (例：図 3)。そのうちいくつかは、ヒト、マウスで報告されている転写開始点と一致したこと、あるいはその 5' 上流域 1-3kbp を「新たな転写調節領域候補」として、*in silico* 解析により転写調節エレメントが抽出されたことから、本来の転写開始点と思われる。

図3: 上流転写開始点の例

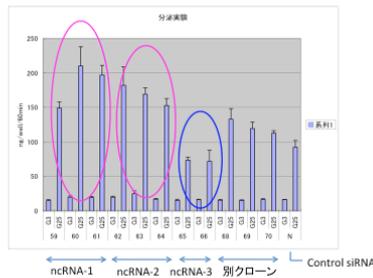


- (3) 新規クローンの機能解析

既報の遺伝子と相関性をもたない新規クローンのうち、500bp 以上のインサートを持ちながら明らかな ORF (open reading frame) を持たない「non-coding RNA(ncRNA)」候補に注目し、このうち 3 クローンについてさらに系統的な解析を進めた。

- ① 臓器分布については、RT-PCR でみる限り比較的 ubiquitous に分布するもの、肝に特徴的に高発現するものを認めた。
- ② 次にグルコース濃度 100mg/dℓの培地と 200 mg/dℓの培地で INS-1 細胞を培養すると、うち 1 クロームは、200mg/dℓグルコースで明らかに発現が低下した。
- ③ 配列特異的な siRNA を INS-1D 細胞へ導入してこの ncRNA 候補の発現を抑制すると、siRNA 導入後 72 時間で、グルコース反応性インスリン分泌 (GSIS) に変

図4:ncRNA候補がGSISへ及ぼす効果:siRNAによる検討



化がみられたものもあった (図 4)。

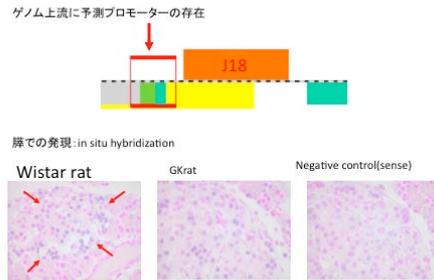
ノックダウンにより分泌亢進、あるいは低下がみられたクローンについては、もとの機能的 ncRNA 候補はそれぞれ何らかの機序で、GSIS に抑制的、あるいは促進的にはたらいていると考えている。

- ④ 機能的 ncRNA 候補の作用機序を探るため、ノックダウンした INS-1D 細胞による網羅的な遺伝子発現プロファイルへの影響をマイクロアレイにより検討した。変動した遺伝子群は、それぞれ 200 をこえたが、特定の既知のパスウェイやネットワークに enrich されてはいなかった。またグルコース反応性インスリン分泌に関わる既報の分子も変動したプロンプの中特に含まれなかったことから、何らかの新規機能的パスウェイの存在が示唆された。これらの変動遺伝子はノックダウンした ncRNA と配列上の相補性をもたず、センサーアンチセンス鎖による調節以外の機序が考えられた。
- ⑤ これらのクローンについて、ラット膵における発現を in situ hybridization (ISH) で検討したところ、プローブ特異的は、発現が認められ、膵組織内では膵ラ氏島に比較的特異的に発現した。また、Wistar ラットと糖尿病 GK ラットとの間で発現に差が認められたクローンが含まれた (図 5)。

(4) 考察

- ① 機能分化した各臓器における「RNA world」の全貌を知り、RNA 発現・調節

図5:新規ncRNA候補の解析(クローンJ18)



機能を解明する方法として、「完全長 cDNA ライブラリー」は、有力な手段であった。特に転写開始点 (TSS) を得ることにより、ゲノム上の転写調節 (プロモーター) 領域を探索できること、RNA の「非翻訳領域」(特に 5' -非翻訳領域) の配列情報が得られること、非コード RNA (ncRNA) の同定ができること、などがあげられる。

「完全長 cDNA」を得る方法もいくつか報告されている。その多くは、哺乳類の mRNA に特徴的な構造 (5' 側のキャップ構造、および 3' 側の poly (A) 付加) を利用するものである。特に日本で菅野純夫博士らにより開発されたオリゴキャップ法は、ヒト全長 cDNA カタログやヒト遺伝子アノテーションデータベース (H-investigational) の構築に多大な貢献をした。

今回用いた Vector-capping 法は、日本で加藤誠志博士により開発された手法であるが、他のライブラリー作製法と比較して、以下のような特徴をもつ。

1. 作製行程で PCR による増幅や制御酵素処理のステップを経ないため、増幅によるバイアスや PCR エラー、微小な欠失などが少ない。
2. ごく比較的少量の RNA (数 μ g) からスタートできる
3. 「完全長」を得る率が高く、しかも 5' 末端配列とゲノム配列との比較から、完全長か否かの判別がある程度可能である。

これらの特徴は、TSS 解析に適しているだけでなく、特に PCR エラーなどがおこりにくいことは、ORF の解析の信頼性が高く、ncRNA の同定に非常に有利であるといえる。

- ② 完全長 cDNA ライブラリーを用いた手法には、コロニーピッキングを含め、自動化していないと大規模な解析に時間と労力が非常にかかる。次に 5' 側からのショットガンシーケンスでは、クローン数が発現量に依存するため、いわゆる「normalization」を行っていないと、発現量の多いクローンほど出現頻度が高くなる。特に多くの ncRNA は、構造タ

ンパクをコードする mRNA に比べて発現量が低いとされており、構造遺伝子の mRNA に比べ確率的に得られにくい、という可能性がある。

今回の cDNA ライブラリーは、5' 例のキャップ構造を 3' 例の poly A 構造を手がかりに「全長」をキャプチャーにしておき、こうした修飾を受けていない RNA は得ることはできない。長鎖 ncRNA は、大きく分けて、遺伝子領域から転写されて、構造タンパク遺伝子の転写制御に関わるもの(アンチセンス鎖など)と、ゲノムアノテーションのないいわゆる intergenic な領域から転写されるものがある。後者の RNA にはポリ A 鎖のついていないものが多いとの報告もあったが、ヒストン修飾に関わるような large intergenic ncRNA (linc RNA) は、mRNA 様の修飾を受けているとされている。我々の解析でも、いわゆる lincRNA とされるクローンが得られている。

- ③ 我々の得た ncRNA は、機能をもつことが予想されるが、ノックダウン後のマイクロアレイの結果をみても、どのようなパスウェイに関連しているのか特定出来ず、そのメカニズムは不明である。

当初 ncRNA の作用機序としては、構造タンパク遺伝子の mRNA と相補的なストランドから転写され、「センス-アンチセンス」のペアにより遺伝子発現を抑制する例が目撃された。しかし ncRNA としてこのような例は必ずしも多くないことが明らかになってきた。また imprinting など広い範囲のゲノム領域の制御に ncRNA がはたらく例も知られている。最近では、ncRNA は RNA に直接作用するのではなく、転写・翻訳の分子マシナリーの一員となり、あるいはその調節因子として作用する例が報告され、注目されている。特にヒストン修飾に関わる linc RNA などは、エピゲノム制御における多彩な役割が明らかになりつつあり、代謝・栄養ほか、細胞の状態や外部因子にもよる細胞機能調節に大きな役割をもつと予想されている。我々の得たクローンのなかにも、グルコース濃度や糖尿病状態で発現が変化する分子もあり、膵β細胞という臓器において、in vivo で生理的・病的意義をもつ可能性があり興味深い。現在、これらを膵β細胞に強制発現させるトランスジェニックマウスの作製などを準備している。

- ④ 本研究開始時に比べ、ここ 2~3 年、こうした分野の研究は飛躍的に発展してきた。その最大の原動力は、次世代シーケンサーの導入である。これは、短い配列断片の情報を並列にして大量に得

るものであり、網羅的な解析に画期的な成果を挙げ、様々な分野で生命科学の概念を根本から変えつつある。

次世代シーケンサーを用いて RNA を解析する (RNA-Seq) 場合、十分なリード数が得られなければ、その転写開始点と転写終始点を必ずしも明確に決定することはできない可能性もある。最近では大量のシーケンスデータを得ることや、第 3 世代以降のシーケンサーの登場により、これらの問題点を克服出来るようになりつつある。

我々も同じ INS-1D 細胞の全 RNA を用いて、イルミナ社 GAIIX を用いて、いわゆる「RNA-Seq」を施行しており、現在解析中である。今後、完全長 cDNA ライブラリーとのデータを比較し、統合する形で解析を進めてゆきたいと考えている。

- ⑤ 今回の研究では以下のことを予定しながら十分行うことができなかったため、今後の課題と考えられる。まず、新規なクローンの全長シーケンスの決定をすすめること。膵β細胞特異的と考えられた転写調節領域について in vitro のレポーターアッセイにより活性を証明し、内在性の DNA 結合分子の単離を試みる。また、転写調節領域の下流にレポーターをつないだトランスジェニックマウスを作成し、組織分布や発現制御機構など、in vivo での意義を探ること。新規 ncRNA については、インスリンプロモータまたは Pdx1 プロモータ下で膵島あるいは膵特異的に発現させたトランスジェニックマウスを作成し、耐糖能異常の有無など表現型を検討すること。これらが挙げられる。

今回得られた新規クローンや、転写調節領域について、ヒトゲノムで対応する分子や領域が得られた場合、多因子遺伝病としての 2 型糖尿病の遺伝素因の、非常に有望な新たな候補領域と考えられる。この観点からも、今後ヒト DNA パネルを用いた解析もすすめてゆきたい。

- (5) [結語]: 我々の作製した全長 cDNA ライブラリーは、膵β細胞における「RNA world」の解明、ひいては糖尿病・代謝疾患の病態の解明に有用なツールと考えられる。

[謝辞]

完全長 cDNA ライブラリーの作製やシーケンスの取得、クラスタリング解析、などにおいて、日立製作所 梶陽介博士、日立ハイテク マニファクチャ&サービス 大島健志朗博士をはじめとする両社の方々にも多大な御

世話になりました。ここに深く感謝の念を記します。

5. 主な発表論文等

〔雑誌論文〕(計 1 件)

安田和基「Regulatory RNA と糖尿病」最新医学 67 巻 1 号 2012

〔学会発表〕(計 1 件)

安田和基「遺伝因子からみた 2 型糖尿病の成因」第 28 回日本医学会総会 平成 23 年 4 月
(東日本大震災のためデータ登録)

〔図書〕(計 1 件)

南茂隆生、安田和基「Non-coding RNA による代謝調節」糖尿病学イラストレイテッド 羊土社 2012

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕なし

6. 研究組織

(1) 研究代表者

国立国際医療研究センター

・研究所・部長

安田和基 (KAZUKI YASUDA)

研究者番号：80311611

(2) 研究分担者

国立国際医療研究センター

・研究所・室長

岡村匡史 (TADASHI OKAMURA)

研究者番号：00333790

(3) 連携研究者 なし