

201135014A

厚生労働科学研究費補助金

難病・がん等の疾患分野の医療の実用化研究事業

(難病関係研究分野)

全ゲノムエクソン配列解析法による

先天性内分泌疾患の分子基盤の解明

(H23-実用化(難病)-一般-014)

平成23年度総括・分担研究報告書

研究代表者 長谷川 奉延

平成24年5月

平成 23 年度研究班構成員名簿

区分	氏名	所属等	職名
研究代表者	長谷川 奉延	慶應義塾大学医学部小児科学教室	准教授
研究分担者	石井 智弘	慶應義塾大学医学部小児科学教室	助教
研究分担者	鳴海 覚志	慶應義塾大学医学部小児科学教室	助教
研究分担者	古川 徹	東京女子医科大学	教授
研究分担者	安達 昌功	神奈川県立こども医療センター内分泌代謝科	部長
研究分担者	長谷川 行洋	東京都立小児総合医療センター内分泌代謝科	部長
研究分担者	荒木 俊介	産業医科大学医学部小児科	助教

目次

I.	総括研究報告書	1
	長谷川奉延 慶應義塾大学医学部小児科学教室・准教授	
II.	分担研究報告書	
1.	先天性内分泌疾患ゲノム DNA 検体収集のための取り組み	3
	鳴海覚志 慶應義塾大学医学部小児科学教室・助教	
	安達昌功 神奈川県立こども医療センター内分泌代謝科・部長	
	長谷川行洋 東京都立小児総合医療センター内分泌代謝科・部長	
	荒木俊介 産業医科大学医学部小児科・助教	
2.	ベンチトップ型次世代シーケンサーを用いた高速・包括的な先天性内分泌疾患遺伝子解析系の確立	5
	長谷川奉延 慶應義塾大学医学部小児科学教室・准教授	
	鳴海覚志 慶應義塾大学医学部小児科学教室・助教	
3.	ヒト 15 器官の RNA 発現データベースの作成	9
	石井智弘 慶應義塾大学医学部小児科学教室・助教	
	鳴海覚志 慶應義塾大学医学部小児科学教室・助教	
4.	全ゲノムエクソン配列解析法による先天性内分泌疾患の分子基盤の解明	13
	古川徹 東京女子医科大学・教授	
III.	研究成果の刊行に関する一覧表	15
IV.	研究成果の刊行物・別冊	17

I. 総括研究報告書

全ゲノムエクソン配列解析法による先天性内分泌疾患の分子基盤の解明

研究代表者 長谷川 奉延 慶應義塾大学医学部小児科学教室 准教授

研究要旨

1. 先天性内分泌疾患患者に由来するゲノム DNA 検体を全国規模で収集するシステムを確立した。
2. ベンチトップ型次世代シーケンサーを用いた高速・包括的な先天性内分泌疾患遺伝子解析系「Endocrinome v1」を世界に先駆けて確立した。
3. 内分泌器官（甲状腺、精巣、卵巣）を含むヒト正常 15 器官の RNA 発現量データベースを作成した。
4. 先天性内分泌疾患患者において次世代シーケンサーを用いた標的領域リシーケンスを行い、多数の一塩基置換、微小挿入欠失を見出した。

A. 研究目的

本研究の目的は、全ゲノムエクソン配列解析法を用いて先天性内分泌疾患の新規責任遺伝子を同定し、本疾患の分子病態を解明することである。その必要性は、本疾患が低身長、肥満、知能障害、不完全な性成熟、不妊などを介して生涯にわたる QOL 低下と後遺症を残す難治性疾患であり、患者数が少なく希少性の高い疾患であるためである。

平成 23 年度（初年度）の目標は、既知遺伝子変異陰性の家系例、孤発例で全ゲノムエクソン配列解析を行うことである。次年度以降、さらに大規模に既知遺伝子変異陰性の家系例、孤発例の解析を進め、候補遺伝子と疾患との関連を、検証コホートでの変異検索、および分子機能解析 で検証する予定である。

B. 研究方法

既知遺伝子変異陰性の家系例、孤発例で全ゲノムエクソン配列解析を行うために、以下のように方法を構築した。

1. 日本小児内分泌学会の承認のもと全国規模での検体収集を行う。
2. 先天性内分泌疾患の発症に関わる多数の遺伝子の塩基配列を高速解析するシステムを確立する。本システムは先天性内分泌疾患の既知責任遺伝子の高速変異スクリーニングに適用

できる。

3. 内分泌器官を含むヒト 15 器官の RNA 発現量のデータベースを作成する。先天性内分泌疾患の新規責任遺伝子同定を目指す本研究において、内分泌器官に特異的に発現する遺伝子のカタログ作成は効率的に研究を遂行する上で有用と考えられるからである。
4. 先天性内分泌疾患患者において、全ゲノムエクソン配列解析を行い、変異を同定する。

（倫理面への配慮）

本研究は慶應義塾大学医学部倫理委員会の承認を得た。また血液採取、遺伝子解析に関しては書面による同意を本人あるいは家族から得てから行った。

C. 研究結果

1. 2011 年 12 月初旬に、日本小児内分泌学会（理事長 横谷 進）の承認のもと、全国の日本小児内分泌学会評議員宛に「『全ゲノムエクソン配列解析法による先天性内分泌疾患の分子基盤の解明研究』参加のお願い」の書面を郵送した。また多数の先天性内分泌疾患患者を継続診療中と考えられる全国 20 施設の部門責任医師に対し、2011 年 12 月中旬に E-mail での研究参加依頼を行った。この結果 7 検体が新たに収集され、現在もさらなる検体が集まりつ

つある。

2. 先天性内分泌疾患の発症に関わる 94 標的遺伝子の全領域を解析可能とした。

3. RNA-シーケンスのデータを利用し、甲状腺、精巣、卵巣を含むヒト 15 器官の RNA 発現量のデータベースを作成した。

4. 先天性内分泌疾患患者における標的領域リシーケンスを行い、多数の一塩基置換、微小挿入欠失を見出した。

D. 考察

1. 今回、全国規模で研究協力を呼びかけることにより、短期間で新規 7 検体を収集することができた。今後も全国規模で積極的な検体収集を継続し、効果的に研究を推進してゆく予定である。

2. 本研究では、先天性内分泌疾患・周辺疾患に関わる 97 遺伝子を高速解析する遺伝子解析システム Endocrinome v1 を確立した。これは世界初の先天性内分泌疾患を対象とした高速かつ包括的な遺伝子解析系である。

3. RNA-シーケンスのデータを利用し、甲状腺、精巣、卵巣を含むヒト 15 器官の RNA 発現量のデータベースを作成した。本研究で用いた RNA-シーケンスのデータは、アッセイ間変動が極めて小さく、良質と考えられた。各器官の特異的遺伝子が高い selectivity index を示しており、ウェット実験、ドライ解析とも有効に行われたことが裏付けられた。

4. 標的領域リシーケンスによる配列解析ではエクソン及び近傍のイントロン部分の塩基配列の完全解明が可能であった。見出されたバリエーションと疾患との関連解析に極めて有用と考えられる。

E. 結論

平成 23 年度は、1. 先天性内分泌疾患患者に由来するゲノム DNA 検体を全国規模で収集するシステムの確立、2. ベンチトップ型次世代シーケンサーを用いた高速・包括的な先天性内分泌疾患遺伝子解析系「Endocrinome v1」の確立、3. 内分泌器官（甲状腺、精巣、卵巣）を含むヒト正常 15 器官

の RNA 発現量データベースの作成、4. 次世代シーケンサーを用いた標的領域リシーケンスによる多数の一塩基置換、微小挿入欠失の同定、において大きな成果を得た。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

Masaki Takagi, Tomohiro Ishii, Aileen M Barnes, MaryAnn Weis, Naoko Amano, Mamoru Tanaka, Ryuji Fukuzawa, Gen Nishimura, David R Eyre, Joan C Marini and Tomonobu Hasegawa. A novel mutation in *LEPRE1* that eliminates only the KDEL ER- retrieval sequence causes non-lethal osteogenesis imperfect. PLoS One. 2012 (in press)

2. 学会発表

該当なし。

H. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

先天性内分泌疾患の遺伝子変異検出法（出願予定）

2. 実用新案登録

該当なし。

3. その他

該当なし。

II. 分担研究報告書

先天性内分泌疾患ゲノム DNA 検体収集のための取り組み

研究分担者 鳴海 覚志 慶應義塾大学医学部小児科学教室 助教
研究分担者 安達 昌功 神奈川県立こども医療センター内分泌代謝科 部長
研究分担者 長谷川 行洋 東京都立小児総合医療センター内分泌代謝科 部長
研究分担者 荒木 俊介 産業医科大学医学部小児科 助教

研究要旨

先天性内分泌疾患は下垂体、甲状腺、性腺など内分泌器官の先天異常の総称であり、いずれの疾患も稀少疾患である。先天性内分泌疾患の分子遺伝学的基盤の解明には多数のゲノム DNA 検体を収集し解析することが必須である。我々は、できるだけ多くのゲノム DNA 検体を確保するために、日本小児内分泌学会の承認のもと全国規模での検体収集を試みた。

A. 研究目的

先天性内分泌疾患患者に由来するゲノム DNA 検体を全国規模で収集する。

考えられる全国 20 施設の部門責任医師に対し、2011 年 12 月中旬に E-mail での研究参加依頼を行った。

B. 研究方法

（1）書面による研究協力の依頼

2011 年 12 月初旬に、日本小児内分泌学会（理事長 横谷 進）の承認のもと、全国の日本小児内分泌学会評議員宛に「『全ゲノムエクソン配列解析法による先天性内分泌疾患の分子基盤の解明研究』参加のお願い」の書面を郵送した。

本書面には ① 解析対象とする疾患の名称（先天性下垂体機能低下症、先天性甲状腺機能低下症、先天性ステロイドホルモン産生異常、先天性カルシウム代謝異常） ② 研究協力施設の役割（インフォームドコンセント、末梢血採取と送付、臨床情報の提供） ③ 研究班の役割（既知責任遺伝子の変異解析、エクソーム解析、結果報告） ④ 費用負担 ⑤ 成果発表 ⑥ 対象基準の詳細（家族例陽性例、重症例、臨床的特徴のある例） ⑦ 試料採取プロトコール について記載した。研究協力施設と研究班との役割分担について、図 1 に示す。

（2）E-mail による研究協力の依頼

多数の先天性内分泌疾患患者を継続診療中と

（倫理面への配慮）

検体収集は、ヒトゲノム・遺伝子解析研究に関する倫理指針に準拠して行われた。研究参加への同意は書面を用いて全ての患者（もしくは患者の親権者）から得た。本研究は、慶應義塾大学医学部倫理委員会で承認されている（承認期間：平成 23 年 11 月 28 日～平成 28 年 3 月 31 日）。

C. 研究結果

書面での研究協力依頼に応じる形で検体送付を行った施設は 2 施設（患者 4 名）であった。E-mail で研究協力依頼を行った 20 施設中 19 施設から返答があり、いずれも研究協力の意向を示すものであった。このうち 3 施設（患者 3 名）から、検体送付があった。

送付検体の疾患別内訳は、先天性下垂体機能低下症 2 名、先天性副腎機能低下症 3 名、先天性カルシウム代謝異常症 1 名、先天性甲状腺機能低下症 1 名であった。

D. 考察

先天性内分泌疾患は下垂体、甲状腺、性腺など内分泌器官の先天異常の総称である。このうち、マススクリーニングで軽症例も含めて診断される先天性甲状腺機能低下症の有病率が 1/4,000 と最も高いが、他疾患の有病率は 1/10,000 未満であり、いずれも稀少疾患である。本研究は、先天性内分泌疾患に関わる未知責任遺伝子の同定を通じて疾病成立機序解明を目指すものであるが、この目的の達成には多数の患者由来ゲノム DNA 検体の収集が不可欠である。

これまで、慶應義塾大学小児科では 500 名以上の先天性内分泌疾患患者の遺伝子解析を行ってきたが、今回、全国規模で研究協力を呼びかけることにより、短期間で新規 7 検体を収集することができた。これらはいずれも家族歴陽性、重症例、臨床的特徴のある例など研究的価値の高い臨床検体であった。今後も全国規模で積極的な検体収集を継続し、効果的に研究を推進してゆく予定である。

E. 結論

書面および E-mail を用いて先天性内分泌疾患ゲノム DNA 検体の収集を試み、新たに 7 名分の検体を収集した。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

該当なし。

2. 学会発表

該当なし。

H. 知的財産権の出願・登録状況

（予定を含む。）

1. 特許取得

該当なし。

2. 実用新案登録

該当なし。

3. その他

該当なし。

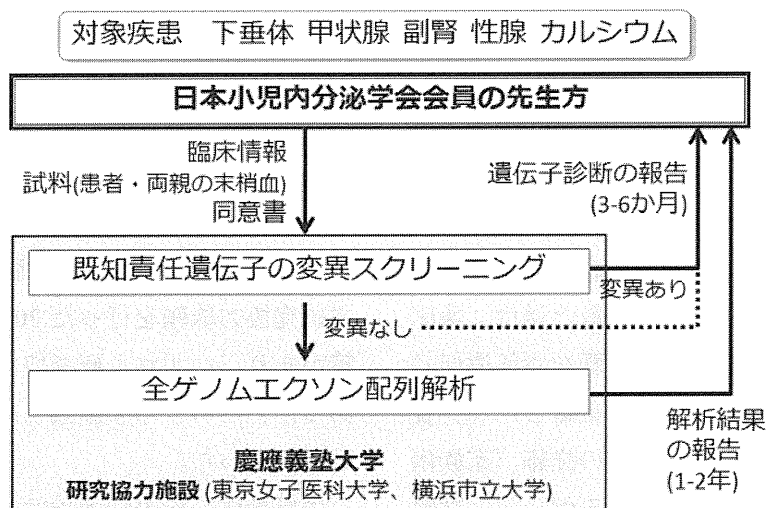


図 1 研究協力施設と研究班との役割分担

ベンチトップ型次世代シーケンサーを用いた 高速・包括的な先天性内分泌疾患遺伝子解析系の確立

研究代表者 長谷川 奉延 慶應義塾大学医学部小児科学教室 准教授

研究分担者 鳴海 覚志 慶應義塾大学医学部小児科学教室 助教

研究要旨

先天性内分泌疾患の発症に関わる 97 遺伝子の塩基配列を高速解析するシステム「Endocrinome v1」を世界に先駆けて確立した。本システムは先天性内分泌疾患の既知責任遺伝子の高速変異スクリーニングに適用できる。また、同様の手法を応用することにより、研究期間中に同定が期待される新規責任遺伝子のスクリーニングも可能である。

A. 研究目的

先天性内分泌疾患・周辺疾患に関わる 97 遺伝子を高速解析する「Endocrinome v1」システムの確立。

B. 研究方法

（1）試料

慶應義塾大学医学部小児科で保管する、既に遺伝子診断の確定した先天性内分泌疾患患者由来ゲノム DNA 検体 12 例分を試料とした。疾患内訳は、先天性甲状腺機能低下症 10 例、先天性副腎機能低下症 2 例。

（倫理面への配慮）

本研究は、ヒトゲノム・遺伝子解析研究に関する倫理指針に準拠して行われた。遺伝子解析研究への参加に関する書面での同意を全ての患者（もしくは患者の親権者）から取得した。

（2）解析対象遺伝子

解析対象とする先天性内分泌疾患および周辺疾患として、先天性下垂体機能低下症、先天性甲状腺機能低下症、先天性副腎機能低下症、性分化疾患、高インスリン性低血糖症、先天性糖尿病、先天性カルシウム代謝異常症、骨形成不全症を選択した（対象遺伝子のリストを表 1 に示す）。

（3）遺伝子解析

Agilent 社のウェブページ eArray (<https://earray.chem.agilent.com/earray/>) を利用して表 1 に示す遺伝子のエクソン領域を選択的に捕捉するための相補的 RNA 群（ベイト）をデザインした。このベイトを含む、Agilent 社の標的 DNA 配列濃縮キット（SureSelect XT Custom Kit）を用いて、試料の調製を行った。

調製したサンプルは 12 検体分をプールし、ベンチトップ型次世代シーケンサー MiSeq で解析した（150 塩基、ペアエンドモード）。出力された配列データは、Burrows-Wheeler Aligner プログラムを用いて二次解析を、Picard、Genome Analysis Toolkit、snpEff、ANNOVAR を用いて三次解析を、それぞれ行った。得られた変異解析結果は、過去に従来型シーケンサーにより取得済みの解析結果と比較した。

C. 研究結果

Endocrinome v1 システムによる解析では 97 標的遺伝子の平均被覆率は $97 \times \sim 280 \times$ （中央値 $169 \times$ ）と極めて良好であった。しかしながら、GC 含有率の高い領域（図 1）、反復配列のためベイトが作製できない領域（図 2）では、被覆率が低下する傾向が確認された。上記 2 つの理由で全領域を解

析できない遺伝子は 97 遺伝子中 3 遺伝子(3.1%)であった。

次に、Endocrinome v1 システムによる変異解析結果を、従来法の結果と比較した（変異の内訳：DAX1 変異 2 種、PAX8 変異 2 種、TG 変異 2 種、TSHR 変異 2 種、DUOX2 変異 10 種）。従来法で同定されている 12 名 18 変異のうち、17 変異は Endocrinome v1 システムで同定可能であった。Endocrinome v1 システムで同定された変異には、19 塩基の欠失、15 塩基の重複も含まれた。

18 変異のうち唯一同定できなかった変異は、DUOX2 遺伝子のエクソン 6 に存在する変異であった。本変異を有する患者の三次解析において、中間段階で生成されたデータを確認したところ、当該変異は snpEff による解析過程で「信頼性の低い結果」として除去されたことが判明した。解析過程から snpEff を除くことにより、DUOX2 エクソン 6 の変異も検出可能であることが再解析実験で確認できた。一方、snpEff を解析過程から除く措置により、信頼性の低い変異検出も含めて報告されることとなった。

D. 考察

本研究では、先天性内分泌疾患・周辺疾患に関わる 97 遺伝子を高速解析する遺伝子解析システム Endocrinome v1 を確立した。これは世界初の先天性内分泌疾患を対象とした高速かつ包括的な遺伝子解析系である。

Endocrinome v1 では、97 遺伝子中 94 遺伝子で十分な被覆率が得られた。SureSelect XT Custom Kit を利用した標的配列濃縮が有効に機能したことを裏付けるデータと考えられる。一方、少数ながら被覆率が不十分な領域を有する遺伝子が存在した。これら遺伝子の解析には、従来法の併用が必要と考える。

本分担研究では、Endocrinome v1 システムの性能検証として、既知の 18 変異の検出を試み、17 変異（94%）の検出に成功した。特筆すべきは、次世代シーケンサーによる変異解析の弱点とされ

る欠失・挿入変異も同定しえた点である。

本分担研究で当初採用したバイオインフォマテイクス解析（二次解析・三次解析）では、DUOX2 エクソン 6 の変異を同定できなかった。これはパラログ遺伝子 DUOX1 の干渉による結果と考えられた。すなわち、DUOX1 には DUOX2 エクソン 6 と DNA レベルで 99.9%相同である領域が存在するため、DNA 断片化処理を経る次世代シーケンサー解析では DUOX1 と DUOX2 の相同を区別できず、変異アレルの割合が過小評価される（通常ならば野生型アレル 1 本、変異型アレル 1 本であるが、パラログまで含むと野生型アレル 3 本、変異型アレル 1 本となる）ことが要因と考えられた。DUOX2 エクソン 6 の変異も、三次解析アルゴリズムを調整し同定可能であった。この偽陰性シグナルを減ずる調整は、同時に偽陽性を増加させる「副作用」を伴うが、相同領域のない部位に生じた「信頼性の低い結果」は偽陽性とすみやかに判定できるため、アルゴリズム改変に伴う解析結果解釈への負荷は限定的と考えられる。

E. 結論

ベンチトップ型次世代シーケンサーを利用した Endocrinome v1 システムにより、先天性内分泌疾患・周辺疾患を高速かつ包括的に解析することができる。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

該当なし。

2. 学会発表

該当なし。

H. 知的財産権の出願・登録状況

（予定を含む。）

1. 特許取得

先天性内分泌疾患の遺伝子変異検出法（出願予定）

2. 実用新案登録

該当なし。

3. その他

該当なし。

表 1 Endocrinome v1 の解析対象遺伝子

下垂体機能低下症：*CGA, GLI2, HESX1, LHX3, LHX4, OTX2, PAX6, POU1F1, PROP1, SIX6, SOX2*

甲状腺機能低下症：*DUOX2, DUOXA2, FOXE1, IYD, NKX2-1, NKX2-5, PAX8, SLC26A4, SLC5A5, TG, TPO, TRH, TRHR, TSHB, TSHR*

副腎機能低下症：*AAAS, CYP11A1, CYP11B1, CYP11B2, CYP17A1, CYP21A2, HSD3B2, MC2R, MRAP, NROB1, POR, STAR, TBX19*

性分化異常症：*AKR1C2, AR, CYB5A, DHCR7, DHH, GATA4, HSD17B3, LHCGR, MAMLD1, MAP3K1, NR5A1, RSPO1, SOX9, SRD5A2, SRY, TSPYL1, WNT4, WT1, ZFPM2*

カルシウム代謝異常：*AIRE, CASR, GATA3, GCM2, GNAS, PRKARIA, PTH, TBCE*

糖代謝異常：*ABCC8, BLK, GCK, GLUD1, HNF1A, HNF1B, HNF4A, INS, KCNJ11, KLF11, NEUROD1, PAX4, PDX1*

骨系統疾患：*BMP1, COL1A1, COL1A2, CRTAP, FKBP10, FGFR3, LEPRE1, LRP5, PPIB, SERPINF1, SERPINH1, SP7*

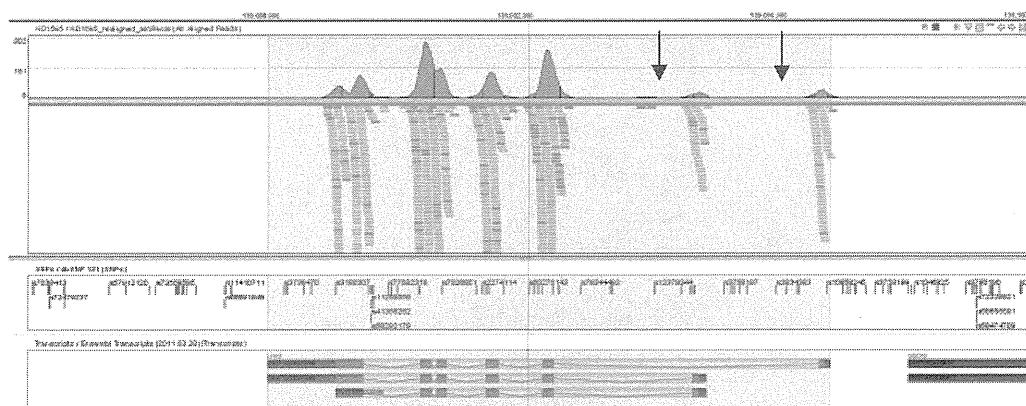


図 1 高 GC 含有率による被覆率の低下 (*LHX3* 遺伝子)
 矢印は被覆率の低いエクソンを示す。

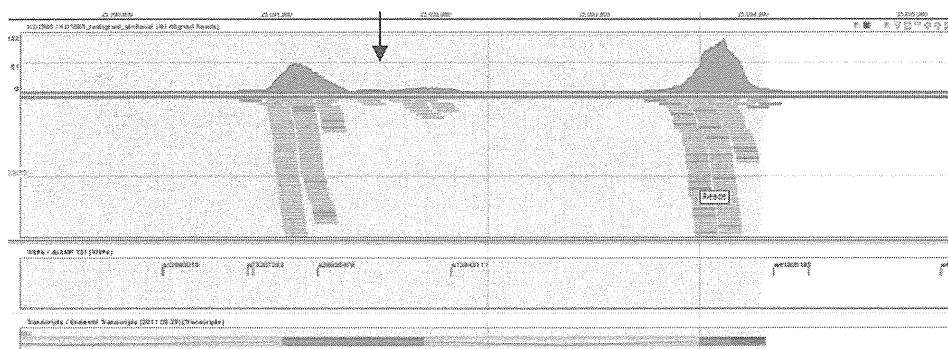


図 2 反復配列による被覆率の低下 (*ARX* 遺伝子)
 矢印は被覆率の低い領域を示す。

ヒト 15 器官の RNA 発現データベースの作成

研究分担者 石井 智弘 慶應義塾大学医学部小児科学教室 助教
研究分担者 鳴海 覚志 慶應義塾大学医学部小児科学教室 助教

研究要旨

先天性内分泌疾患の新規責任遺伝子同定を目指す本研究において、内分泌器官に特異的に発現する遺伝子のカタログ作成は効率的に研究を遂行する上で有用と考えられる。本分担研究では、公共で利用可能な RNA-シーケンスのデータを利用し、甲状腺、精巣、卵巣を含むヒト 15 器官の RNA 発現量のデータベースを作成した。

A. 研究目的

内分泌器官（甲状腺、副腎、精巣、卵巣）を含むヒト正常器官の RNA 発現量データベースの作成。

B. 研究方法

（1）配列データ

RNA-シーケンスのデータは Harvard 大学が行った健全なヒト 16 器官の RNA-シーケンス解析研究 bodymap 2.0 のものを利用した。bodymap 2.0 では、ヒト献体に由来する 16 器官（副腎、脂肪組織、脳、乳腺、結腸、心、腎、肝、肺、リンパ節、卵巣、前立腺、骨格筋、精巣、甲状腺、白血球。）に由来する RNA を、次世代シーケンサー HiSeq2000 で解析している。各器官とも、同一 RNA を用いて独立した 2 アッセイ（50 塩基ペアエンド、75 塩基シングルエンド）が行われた。配列データは European Bioinformatics Institute の公共サーバーで一般公開されている。本分担研究では、下記ウェブサイトからそれぞれ 16 器官、計 32 アッセイ分のデータをダウンロードした。

<http://www.ebi.ac.uk/arrayexpress/browse.html>

（2）データ解析

取得した配列データの二次解析には、イントロン・ギャップを考慮したアライメントが可能な Linux ベースのプログラム Tophat-1.4.0 を用いた。

定量解析には、Linux ベースのプログラム Cufflinks-1.3.0 および Windows ベースのプログラム Avadis NGS-v1.3 を、それぞれを用いて解析した。発現量データは Cufflinks では Fragments Per Kilobase of exon per Million mapped fragments (FPKM) 値を、Avadis NGS では Reads Per Kilobase of exon model per Million mapped reads (RPKM) 値をそれぞれ指標とした。各器官のデータは Quantile 補正を行った上で比較した。

（3）器官別発現データベースの作成と検証

上記解析で得られた器官別、遺伝子別の Quantile 補正済み RPKM 値（もしくは FPKM 値）を Excel 2007/2010 のシートに展開した。発現量の器官特異性の指標として以下の通り selectivity index を定義した。

$$\text{器官 } O \text{ での遺伝子 } G \text{ の Selectivity index} = \frac{(\text{O での } G \text{ の RPKM 値}) \text{ の平均}}{(\text{O を除く 15 器官での } G \text{ の RPKM 値}) \text{ の平均}}$$

*RPKM 値は FPKM 値で置換可

本データベースの有効性を評価するため、16 器官に特異的に発現する遺伝子の Selectivity index を確認した。

C. 研究結果

本解析により、16 器官につき 19,613 種の遺伝子の RNA 発現量（FPKM 値もしくは RPKM 値）が含まれるデータベースを作成した（図 1）。本データベースは Excel 2007/2010 で容易に扱うことができ、数値にもとづくフィルタリングや統計操作（平均値や中央値の演算処理など）を実行できることを確認した。

Cufflinks で得た FPKM 値と Avadis NGS で得た RPKM 値の間には良好な相関があり、三次解析の出力データにおけるプログラム依存性は小さいと考えられた。以後のデータ検証は RPKM 値にもとづき行った。

bodymap 2.0 のデータは同一検体 2 アッセイ分が含まれるため、まずアッセイ間変動を検証した。その結果、RPKM 値が低い（RNA 発現量の低い）遺伝子ほどアッセイ間変動が大きいことが明らかになった（図 2）。同様の傾向は DNA チップを用いた RNA 発現解析でも指摘されており、RNA 発現解析に共通の現象と考えられた。本研究では RPKM 値 10 未満のデータは再現性が担保されないと判断し、selectivity index 算出のから除外した。

次に、各器官で特異的発現を見せる既知遺伝子（甲状腺での TSH 受容体、肝での C reactive protein など）の selectivity index を参照し、各器官の解析の妥当性を検証した。その結果、副腎を除く 15 器官では、特異的発現の知られる既知遺伝子がおおむね高い selectivity index を示すことが確認できた。副腎では ACTH 受容体、21 水酸化酵素など副腎に特異的に発現する遺伝子の selectivity index および生データ（RPKM 値、FPKM 値）のいずれも低値であり、妥当性が確認できなかった。そのため、副腎に由来するデータは破棄し、計 15 器官として各器官各遺伝子の selectivity index を再計算した。副腎データ除外に伴う selectivity index への影響は 99%以上の遺伝子で 1%未満であった。

D. 考察

本分担研究では RNA-シーケンスのデータを利用し、甲状腺、精巣、卵巣を含むヒト 15 器官の RNA 発現量のデータベースを作成した。本研究で用いた RNA-シーケンスのデータは、アッセイ間変動が極めて小さく、良質と考えられた。16 器官中 15 器官では、各器官の特異的遺伝子が高い selectivity index を示しており、ウェット実験、ドライ解析とも有効に行われたことが裏付けられた。一方、副腎のみ解析の有効性が確認できなかった。① 器官の死後阻血への抵抗性の差異 ② 献体提供者死亡から RNA 抽出処理までの期間の差異 ③ RNA 保管状態の差異 などの影響が推測されるが、原因は不明である。

エクソーム解析では約 2 万種の遺伝子を解析するため、偽陽性の制御、すなわち変異が発見された遺伝子が、真の疾患責任遺伝子か否かを見極めることが肝要である。これまで同定されてきた先天性内分泌疾患の責任遺伝子は、当該内分泌器官に特異的に発現する遺伝子が大半である。このため、今回作成した甲状腺、精巣、卵巣の発現データベースは、それぞれの器官の候補遺伝子の病因としての確からしさを判断する上で極めて重要なリソースとなる。一方、副腎、下垂体など他の内分泌器官の RNA 発現データベースは独自に RNA-シーケンス解析を行う必要があり、今後の課題である。

E. 結論

公共で利用可能な RNA-シーケンスのデータを利用し、甲状腺、精巣、卵巣を含むヒト 15 器官の RNA 発現量のデータベースを作成した。このデータベースは、先天性内分泌疾患の新規責任遺伝子を同定する本研究課題の遂行上、有用と考えられる。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

該当なし。

2. 学会発表

該当なし。

H. 知的財産権の出願・登録状況

（予定を含む。）

1. 特許取得

該当なし。

2. 実用新案登録

該当なし。

3. その他

該当なし。

Gene	thyroid	adipose	brain	breast1	breast2	colon	colon2	heart1	heart2	kidney	kidney2	liver1	liver2	lung1	lung2	lymph	stomach	ovary				
A1GC	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0				
A1CF	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0				
A20I	87.1	77.6	82.7	80.2	28.6	49.8	83.6	87.1	28.0	36.4	83.5	73.0	117.3	127.8	382.8	365.8	9.1	9.4	32.7	20.9	29.1	
A2M1	11822	23697.1	24081.0	25128.5	81176.4	10246.6	10204.0	69434.4	84175.3	97176.4	300310.4	73615.0	73764.0	80963.4	141979.5	117396.0	167203.4	231404.1	105631.0	107948.7	36399.1	
A2M1	11810	39.1	6.9	11.5	84.8	271.8	177.4	37.2	17.2	10.2	5.8	28.1	8.0	23.7	3.4	8.5	0.0	12.1	6.1	22.4	14.1	19.1
AASALT	18708	565.9	515.9	2304.3	1955.2	112.9	112.9	890.9	319.9	1896.8	1853.5	1534.7	1248.3	1810.3	1469.1	151.0	126.7	231.8	1171.4	1349.9	1194.5	642.1
AASNT	4120	0.9	1.7	4.5	4.4	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
AASZ	12138	914.7	924.1	927.7	874.0	944.8	518.6	844.5	890.9	712.4	953.1	918.1	1095.6	915.2	1291.2	1258.9	735.7	762.9	1345.8	1248.1	1092.1	
AACS	12649	1274.6	1360.5	1097.3	1117.3	1572.0	1560.8	2368.6	2351.0	721.4	712.4	925.5	860.5	1309.0	1348.4	590.7	819.1	558.2	1609.6	744.4	884.4	761.1
AADAC	4071	0.2	0.2	619.3	660.4	0.2	0.2	417.9	405.4	5.3	6.1	7.5	0.7	0.2	0.2	19546.1	19813.5	139.4	123.3	0.2	0.2	0.2
AADACL2	4090	0.2	0.2	3.1	0.2	0.2	0.2	2.7	4.9	33.3	34.5	0.2	0.2	0.2	0.6	1.9	4.6	6.5	27.7	25.3	0.2	
AADACL3	166	0.3	0.4	1.1	1.8	0.5	0.7	28.0	34.7	0.8	1.7	2.0	4.4	0.4	0.7	3.6	7.5	0.6	0.9	1.0	1.3	0.2
AADACL4	165	0.3	0.4	1.1	1.8	0.5	0.7	1.9	2.9	1.5	6.3	2.0	4.1	30.7	24.2	3.6	7.5	0.6	0.8	9.5	11.1	0.2
AADAT	5149	346.3	326.9	116.3	111.7	344.0	575.3	131.2	110.5	120.5	312.3	143.6	195.5	139.9	135.1	1033.1	1244.3	60.4	70.7	51.8	55.1	241.1
AAGAS	15999	2991.9	2965.1	1797.8	1808.9	2324.0	2326.7	1674.5	1710.4	1290.8	152.4	1497.2	1474.5	3126.4	3441.9	2092.0	2079.4	2109.0	2334.6	1494.2	1401.7	2073.1
AAGI	2386	781.5	750.7	688.6	777.3	2432.3	1402.5	291.1	194.4	370.1	254.2	2664.8	1640.9	499.5	491.3	268.8	742.3	262.7	1590.8	346.1	402.9	443.1
AAGP	3139	2587.9	2731.7	3393.0	3559.0	2104.9	2562.6	2749.3	2364.4	3013.4	2841.6	2317.0	3155.3	3489.8	3368.8	2457.7	3359.2	2989.1	3102.5	2957.0	3166.2	2294.1
AAGT	10939	7.1	7.7	0.0	0.0	1.7	0.0	0.0	0.0	1.7	5.5	7.0	12.9	0.0	0.0	4.6	9.0	0.0	0.0	3.9	5.8	0.2
AARS	14790	11007.5	11448.5	8245.8	8482.8	19643.7	19128.4	6559.8	6601.5	6445.6	855.8	4713.9	4854.3	7572.8	7558.6	6982.2	6804.0	5151.3	5160.0	4902.5	5150.6	11040.1
AARS2	6058	348.8	342.9	294.7	186.8	234.4	189.7	227.0	297.9	400.5	265.2	393.4	394.1	471.1	411.7	502.1	454.7	265.4	167.9	244.1	161.4	650.2
AARSD1	15254	1424.8	1624.8	1806.6	261.7	1514.7	1414.3	1144.2	1140.7	1202.5	1341.1	2346.6	2381.6	640.5	607.1	1227.4	1245.2	531.9	516.4	587.0	622.8	1276.1
AASDH	4712	564.9	575.3	294.1	328.1	254.6	376.8	610.7	684.0	352.1	392.8	421.3	427.4	624.8	642.3	334.6	562.3	192.9	134.6	402.8	478.4	824.1
AASDHPT	11475	2217.4	2266.2	1674.1	1791.0	7381.9	7380.8	2195.0	3335.5	1782.1	1782.7	3012.0	3154.5	2024.2	2848.2	2941.2	3216.3	1252.4	1277.8	2189.0	1037.5	3401.1
AASZ	7861	1037.5	1020.8	1593.1	1496.3	1037.0	1008.0	529.1	478.2	706.0	630.8	1972.7	2081.4	2039.1	1878.3	3996.3	3514.1	918.8	845.7	822.1	797.3	1394.1
AATF	15467	2705.1	2727.7	4057.3	4171.4	1980.9	1951.4	2173.8	2246.9	4187.9	4187.9	1548.6	1588.8	2025.3	2060.3	1573.5	1512.9	1670.2	1642.2	1634.3	1643.4	3111.1
AATF	15097	172.6	181.4	989.0	883.3	9470.3	4559.9	49.9	47.9	188.0	182.7	194.7	145.0	23.9	24.4	30.1	50.0	307.5	195.0	206.0	165.6	48.1
ABAT	14543	4114.0	4214.0	272.6	239.1	16176.1	13861.3	242.7	160.5	748.6	757.5	1120.6	1005.7	2311.4	1912.4	32311.9	31410.6	397.8	198.3	656.3	802.5	744.1
ABCA3	9381	4712.7	4732.2	11893.7	12123.3	1431.5	1472.8	1421.0	7589.8	8256.2	1649.8	1674.5	1686.1	1333.6	1922.6	18246.7	4342.6	4377.8	3273.4	3266.4	2931.1	

図1 器官別発現データベース

縦列に遺伝子名、横軸に16器官をとり、各セルにはRPKM値が入力されている。

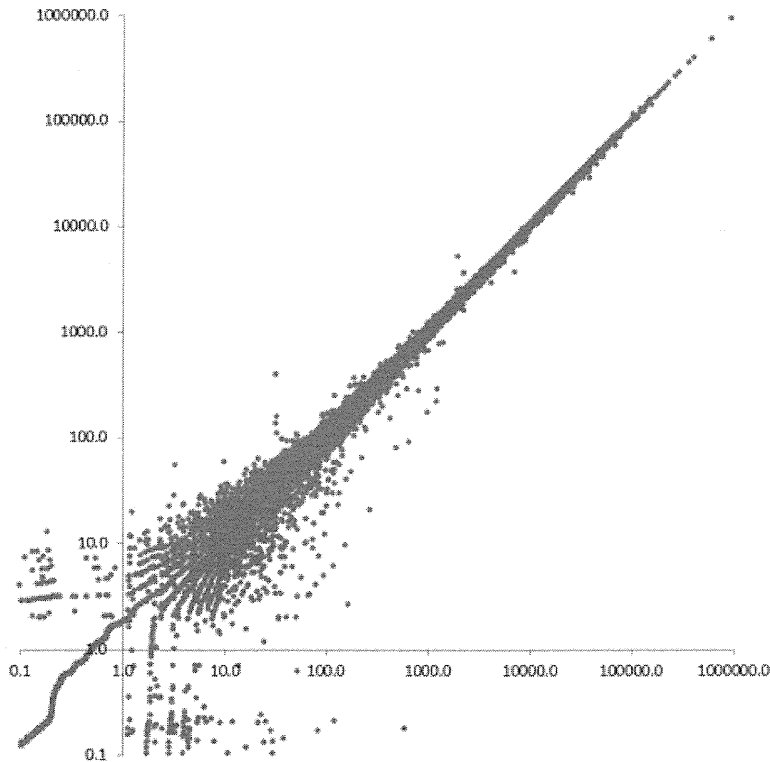


図2 アッセイ間変動（脂肪組織）

横軸はアッセイ1 (75塩基シングルエンド)、縦軸はアッセイ2 (50塩基ペアエンド)から算出されたRPKM値。低発現域はアッセイ間変動が大きくデータの再現性が低い。

全ゲノムエクソン配列解析法による先天性内分泌疾患の分子基盤の解明

研究分担者 古川 徹 東京女子医科大学 教授

研究要旨

先天性内分泌疾患患児における標的領域リシーケンスを行い、多数の一塩基置換、微小挿入欠失を見出した。

A. 研究目的

先天性内分泌疾患はほとんどが遺伝子異常に基づくと考えられるが原因遺伝子については多くが不明のままとなっている。本研究は原因遺伝子不明となっている先天性内分泌疾患の原因遺伝子を比較的広範囲な標的領域リシーケンスあるいは全ゲノムエクソンリシーケンス解析により明らかにすることを目的とする。

B. 研究方法

患児血液より DNA を抽出しライブラリーを作製してゲノムの標的領域あるいは全エクソン部分を抽出し次世代型シーケンサーで塩基配列を解析する。リファレンス配列と比較し、一塩基置換、微小挿入欠失を見出して原因遺伝子を確定する。

（倫理面への配慮）

検体収集は、ヒトゲノム・遺伝子解析研究に関する倫理指針に準拠して行われた。研究参加への同意は書面を用いて全ての患者（もしくは患者の親権者）から得た。本研究は、慶應義塾大学医学部倫理委員会で承認されている（承認期間：平成 23 年 11 月 28 日～平成 28 年 3 月 31 日）。検体をすべて連結可能匿名化し、全ゲノムエクソンリシーケンス解析に供した。

C. 研究結果

先天性内分泌疾患患児における標的領域リシーケンスを行い、多数の一塩基置換、微小挿入欠失を見出した。現在その一塩基置換、微小挿入欠失をリファレンス配列と比較中である。

D. 考察

標的領域リシーケンスによる配列解析ではエクソン及び近傍のイントロン部分の塩基配列の完全解明が可能であり、見出されたバリエーションと疾患との関連解析に極めて有用と考えられる。

E. 結論

標的領域リシーケンスは先天性内分泌疾患の疾患ゲノム解析に有用と考えられた。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

該当なし。

2. 学会発表

該当なし。

H. 知的財産権の出願・登録状況

（予定を含む。）

1. 特許取得

該当なし。

2. 実用新案登録

該当なし。

3. その他

該当なし。

III. 研究成果の刊行に関する一覧表

研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Masaki Takagi, Tomohiro Ishii, Aileen M Barnes, MaryAnn Weis, Naoko Amano, Mamoru Tanaka, Ryuji Fukuzawa, Gen Nishimura, David R Eyre, Joan C Marini and Tomonobu Hasegawa	A novel mutation in <i>LEPRE1</i> that eliminates only the KDEL ER-retrieval sequence causes non-lethal osteogenesis imperfecta	PLos One	in press	in press	2012