

顔面肩甲上腕型筋ジストロフィーの エピジェネティック病態解明と革新的治療法の開発

研究分担者 横田 隆徳 東京医科歯科大学大学院脳神経病態学分野 教授

研究要旨

顔面肩甲上腕型筋ジストロフィー（Facioscapulohumeral muscular dystrophy, FSHD）は、4番染色体テロメア近傍にあるD4Z4反復領域の転写が活性化され、その結果DUX4ホメオボックス転写因子が発現されることが進行性筋萎縮の原因と考えられている。我々は、D4Z4領域の転写活性化にはクロマチン構造制御因子の一つASH1が必要であり、特にD4Z4領域の5'側から転写されるlong non-coding RNA (lncRNA)の一種DBE-TがASH1のSETドメインに結合しASH1をリクルートすることを報告している(Cabianca et al., 2012)。

本研究は、FSHDの分子標的治療薬の開発を目指し、特に我々の研究グループが専門とするASH1とlncRNAとの相互作用に着目してDBE-Tに対するロックダウンヘテロ核酸を合成した。

A. 研究目的

大部分のFSHD患者では4番染色体テロメア近傍にあるD4Z4反復配列が短縮しているが、5'側には全ての患者で残されている領域（NDE; non-deleted element）がある。我々は、NDE領域からnon-coding RNAの一種であるDBE-T（D4Z4-binding element）が転写され、これがASH1のSETドメインに結合することによってASH1をD4Z4領域にリクルートすることを報告している（Cabianca et al., 2012）。

ASH1はHOX遺伝子群の発現維持に関わるTrithoraxグループの一員であり、そのSETドメインによってヒストンH3のリジン36を選択的にメチル化する(Tanaka et al., 2007)。これまでにも他のSETドメインがRNAと結合することは報告された例があるが、HOX遺伝子以外でTrithorax因子がnon-coding RNAと結合することを示したのは、我々の報告が最初である。

D4Z4領域の転写が活性化されると、D4Z4反復配列の最も3'側のD4Z4にコードされるDUX4蛋白質が発現される。DUX4はホメオボックス型の転写因子で、さらに下流のPITX1という別の転写因子の発現を誘導し、その結果筋萎縮を引き起

こすと考えられている。

D4Z4領域の転写活性化からPITX1の発現に至る過程で、FSHDの分子標的治療薬となる可能性のあるDBE-T, ASH1, DUX4, PITX4の4つの標的候補の中で、我々は最も上流に位置するlncRNAがFSHD患者の筋細胞でのみ発現されること（特異性）に注目し、DBE-Tに対するロックダウンヘテロ核酸をデザインした。

様々な人工核酸（オリゴヌクレオチド）によるロックダウン技術の中で、架橋化核酸（Bridged Nucleic Acid）またはLNA（Locked Nucleic Acid）として知られる新しい人工核酸は、二重鎖核酸のリボース立体構造を固定化することにより、標的核酸に対する結合親和性、ヌクレアーゼ耐性、さらに細胞膜透過性を高めるという特徴を有している。さらに、ロックダウンオリゴの構造をRNA-DNA-RNAというヘテロ核酸にすることにより、中央のDNA部分が標的RNAと結合して生じるDNA/RNAヘテロ二重鎖をRNaseHが切断し、それによって標的RNAを効率的に分解することが可能になる（資料4.1A）。

本研究では、約3.4KbとなるDBE-Tに対して20種類のLNA核酸を合成した。

B.研究方法

【LNA gapmer の合成】

ジーンデザイン社 (Eziqon 社) のノックダウンオリゴのデザインアルゴリズムを用いて、DBE-T に対する 20 種類の LNA gapmer (ヘテロ核酸) を合成した。また、DBE-T の発現レベルを定量するための Taqman PCR プローブを合成した。

【DBE-T 発現ベクターの作成】

LNA gapmer の活性を *in vitro* で評価するため、3,374bp の DBE-T を CMV プロモータ下にクローニングした発現ベクターを作成した (資料 4.1B)。

(倫理面への配慮)

本研究は FSHD 患者由来の細胞を使う予定であり、本学医学部倫理審査委員会 (承認番号 199) 及び難治疾患研究所倫理委員会 (承認番号 2014-012) の審査を経て承認を得ている。

C.研究結果

【LNA gapmer の合成】

DBE-T に対する 20 種の LNA gapmer を合成した。

【DBE-T 発現ベクター】

DBE-T は FSHD 患者細胞でのみ発現する。そこで、DBE-T のノックダウン効率をより簡便に定量するために、DBE-T を CMV プロモータ下にクローニングした発現ベクターを作成した。テンプレートとして、PacBio 解析にも用いた BAC クローン PR11-242C3 を用い、シーケンシングによって配列を確認した。DBE-T の一部は GC 含有率が極めて高く、PrimerSTAR GXL DNA ポリメラーゼ (タカラバイオ) によって増幅することを見出すまで時間が掛かってしまった。CMV-DBE-T 発現ベクターを HEK293 細胞に一過性発現させ、RT-PCR

によって DBE-T の発現を確認した。

D.考察

FSHD の分子標的治療薬として、lncRNA の一種である DBE-T に対する LNA gapmer を 20 種合成した。

これらノックダウンオリゴの効果を評価するために、現在 DBE-T 発現ベクターを導入した細胞での Taqman-PCR による定量スクリーニングを行っている。

今後、D4Z4 トランスジェニックマウス由来の線維芽細胞 (MEF) または筋芽細胞でのノックダウン実験、さらに FSHD 患者由来の細胞 (末梢血リンパ球または iPS 細胞由来の分化筋芽細胞) を予定している。

E.結論

DBE-T に対する 20 種のヘテロ核酸 (LNA gapmer) を合成したが、アッセイ系の立ち上げが遅れたため (DBE-T 発現ベクター、D4Z4 トランスジェニックマウス、FSHD 患者細胞) 実際のスクリーニングは現在開始した段階である。

F.健康危険情報

該当なし。

G.研究発表

1. 論文発表
なし
2. 学会発表
なし

H.知的財産権の出願・登録状況

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

顔面肩甲上腕型筋ジストロフィーの エピジェネティック病態解明と革新的治療法の開発

研究分担者 森岡 勝樹 東京医科歯科大学難治疾患研究所声明情報学分野 助教

研究要旨

顔面肩甲上腕型筋ジストロフィー（Faciocapulothoracic muscular dystrophy, FSHD）の原因となる4番染色体テロメア近傍のD4Z4領域は、約3.3Kbの高度に保存されたGCリッチな配列が数十コピー反復するマイクロサテライトリピートの一種である。従来のSangerシーケンス法や次世代シーケンス法では、D4Z4の反復単位よりも短い配列しか得られないため、それぞれの反復単位を正確に並べてアセンブルすることは不可能に近かった。実際、4番染色体のテロメア近傍と、これと相同性の高い領域を含む10番染色体の領域は今日に至るまでシーケンス情報がほとんど無い。

本研究では、FSHDの新たな遺伝子診断技術の確立を目指して、D4Z4領域を含むBACクローンを用いてPacBio RSによるde novoアセンブリを試みた。また、得られたシーケンス情報を用いて、実際にD4Z4領域のde novoアセンブリに必要なリード数や配列条件をシミュレーション実験によって検証した。

A. 研究目的

ヒト4番染色体のテロメア近傍には、3.3Kbのレトロ反復配列（D4Z4）が健常人では約100コピー存在する（資料4.2A）。顔面肩甲上腕型筋ジストロフィー（FSHD）患者では、D4Z4のコピー数が10以下に短縮しているか（1型）、またはヘテロクロマチン制御因子SMCHD1遺伝子に変異している（2型）ことが知られている。その結果、D4Z4のヘテロクロマチン構造が緩み、D4Z4にコードされるホメオボックス転写因子DUX4が発現されることが筋萎縮の原因になると考えられている。患者の筋細胞でD4Z4が転写されてしまうのは、ヒストンメチル基転移酵素ASH1がnon-coding RNAによってD4Z4領域にリクルートされ、D4Z4の転写を活性化することが原因の一つと考えられており（Cabianca et al., Cell, 2012）、エピジェネティックな作用機序に着目した新たな治療法の開発が期待されている。

FSHDは10万人に5～12人の発症頻度と言われているが、その遺伝子診断にはSouthernプロット法やFISH（Molecular combing法）によるD4Z4のコピー数評価が必要となるため、ハードルが高かった。

そこで、これらに代わる診断法として、一分子DNAシーケンサーによるゲノムアセンブリがD4Z4領域において原理的に可能かどうかを、D4Z4を含むBACクローンを用いて検証した。

4番染色体のD4Z4反復配列は、それぞれが高い相同性を持つばかりでなく（資料4.2B）、ほぼ同じ反復配列が10番染色体上にも存在する。従って、3.3Kbのリピート配列を正しくアセンブルするためには、少なくともリピート単位よりも長いリードが必須となり、現状ではPacBio RSに限られる。PacBio RSの長所は、第三世代シーケンサーに比べて得られる配列情報が圧倒的に長く（30Kbを越える）PCRを介さない一分子シーケンサーでGC含有率によるバイアスが全く無いことである。一方In/Delエラーが多い、得られるリード数が少ないといった短所もある。

B. 研究方法

【BACクローンのシーケンシング】

D4Z4領域全体を含む新たなBACクローンは、BACPAC Resources（Children's Hospital Oakland Research Institute, USA）より入手した。このライブラリーのイ

ンサートの長さは平均 80Kb である。20 μ g の DNA を D4Z4 反復の外側で切断する EcoRV で断片化し、濃縮 (AMPure), SMRT bell ライブラリ作成 (Pacific Bioscience) を行い、DNA/Polymerase Binding Kit: P4 または P5、Sequencing reagent: v2.0 (C2 または C3)、SMRT cell: v3 を用いて 1 セルのシーケンスを行った。

得られたシーケンスデータを用いて、PacBio による HGAP 解析パイプラインによる de novo アセンブリを行った。また、約 15 万リードのシーケンスデータの一部 (100~4 万) をランダムに選び、de novo アセンブリが可能かどうか、またどのような配列情報がアセンブリに影響するかをシミュレーション実験によって検証した。

(倫理面への配慮)

この実験は既存の BAC クローンを用いて行うため、倫理面での問題は発生しない。

C. 研究結果

【de novo アセンブリの結果】

同じ BAC クローンを用いて、P4-C2 または P5-C3 の反応ケミストリを用いてシーケンシングを行った。一般的には P5-C3 の方がより長いリードが得られる。シーケンシングの結果、P4-C2 と P5-C3 ケミストリによって得られるリード数、マップされたサブリードの数や精度には大きな違いは無かったが (資料 4.2C)、HBAP2/3 解析パイプラインでは P4-C2 ケミストリのデータのみ de novo アセンブリによって D4Z4 全体をカバーするコンティグが得られた (資料 4.2D)。そこで、P4-C2 と P5-C3 の間でより詳細なリードの比較を行った所、リードの長さの分布には違いが無かったが、P5-C3 ケミストリで得られた配列は特に In/Del エラーが多いことが分かった (資料 4.2F)。

【シーケンスデータを用いたシミュレーション実験】

de novo アセンブリに成功した P4-C3 のデータを用い、ランダムに選んだ 100~4 万リードを用いて de novo アセンブリを行った結果、最小で 1000 リードから 4000 リードあれば、D4Z4 領域のアセンブリが可

能であることが分かった (資料 4.2G)。即ち、PacBio RS の 1 セル当たりのリード数約 15 万の 1/40~1/150 のデータでアセンブリが可能ということになる。さらに、このシミュレーションにおいてうまくアセンブリされたデータセットとアセンブリされなかったデータセットを比較検討したところ (資料 4.2H)、約 15Kb 以上の長さを持つ D4Z4 領域のリードが 1 つ以上含まれることが必須であることが明らかになった。

D. 考察

これまで D4Z4 のような長い反復配列のシーケンシングは極めて困難であり、実際最新のヒトゲノムデータベースにおいても D4Z4 領域は大きなギャップとして残されている。今回、BAC クローンをを用いた検証実験ではあるが、PacBio RS によって 3.3Kb の D4Z4 配列を 13.5 コピー含む DNA を de novo アセンブリ出来ることが実証された意義は大きい。

実際のシーケンスデータを用いて、D4Z4 領域のアセンブリに必要な条件を検証し、うまくアセンブリするためにはリード数は必ずしも多い必要はなく、それよりも D4Z4 反復配列を 4 個以上含む長さのリードが 1 つ以上含まれることが重要であることが示された。FSHD 患者の大部分 (95%) は、D4Z4 が 10 コピー以下に短縮しているため、今回の検証実験の解析条件はそのまま患者の遺伝子診断に適用可能である。一方、健常人ではそれよりも長い反復配列を含むため、今回の BAC クローンよりも厳しい条件 (リード数と長さ) が要求される可能性が高い。

今回の解析では、EcoRV によって切断した DNA 断片をシーケンシング解析に用いている。患者の遺伝子診断では、このような EcoRV に相当するゲノム DNA を抽出する必要がある。そのために、我々は並行する研究で FLAG-dCas9 と D4Z4 特異的なガイド RNA により D4Z4 を含むゲノム DNA 断片を精製する方法を確立しており、今後実際の患者ゲノム DNA を用いた検証を進める予定である。

E. 結論

PacBio RS により、D4Z4 を 10 コピー以

上含む DNA 断片の *de novo* アセンブリに成功した。また、現行の PacBio RS の 1 セル当たり得られるデータの数十分の 1 の少ないデータ量でも、比較的長いリードが含まれればアセンブリ可能であることを実証した。並行する研究で D4Z4 領域を特異的に抽出する技術も確立しており、今後患者ゲノム DNA を用いた遺伝子診断技術の検証を進める予定である。

F.健康危険情報

該当なし。

G.研究発表

1 . 論文発表

Sequencing and *de novo* assembly of macrosatellite repeats of the FSHD locus. Morioka, MS, Tanaka, Y. (submitted)

De novo assembly of PacBio sequence data by homopolymer contraction method. Morioka, MS, Tanaka Y. (manuscript in preparation)

2 . 学会発表

第 37 回日本分子生物学会年会・平成 26 年 12 月 (横浜)・PacBio RS による顔面肩甲上腕型筋ジストロフィー遺伝子座のシーケンシング (田中裕二郎・森岡勝樹)

H.知的財産権の出願・登録状況

1 . 特許取得

学内審査中 (ホモポリマー縮合による *de novo* アセンブリの効率化)

2 . 実用新案登録

なし

3 . その他

なし