

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 5 月 28 日現在

機関番号：12702

研究種目：若手研究(B)

研究期間：2010～2012

課題番号：22770239

研究課題名（和文） ヒトのコーディング領域トリプレットリピートの進化的意義

研究課題名（英文） The evolution of coding triplet repeats in human genome

研究代表者

五條堀 淳 (GOJOBORI JUN)

総合研究大学院大学・先導科学研究科・助教

研究者番号：00506800

研究成果の概要（和文）：コーディング領域トリプレットリピートはタンパク質の構造中、不確定領域を形成する事が知られ、そのような領域は進化速度が速いと言われていた。本研究でトリプレットリピートの進化とタンパク質の構造の関連を調べた結果、霊長類における進化では、これまで言われていた事とは逆の、不確定領域になりにくいトリプレットリピートの進化速度が速い事が示され、霊長類特異的な進化の特徴が明らかにされた。

研究成果の概要（英文）：It is proposed that triplet repeats in coding regions are more likely to form disordered region in protein structure. It is suggested that disordered regions evolve faster. In this study, I examine the relationship between the evolutionary rate and protein structure for triplet repeats in coding regions. I showed that the repeats in non-disordered region evolved faster than disordered region during primate evolution. Here I revealed primate specific evolution on coding triplet repeats.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	900,000	270,000	1,170,000
2011年度	600,000	180,000	780,000
2012年度	500,000	150,000	650,000
年度			
年度			
総計	2,000,000	600,000	2,600,000

研究分野：分子進化学

科研費の分科・細目：生物系・生物学・自然人類学

キーワード：アミノ酸リピート、不確定領域

1. 研究開始当初の背景

- (1) ハンチントン舞踏病、脊髄小脳変性症に代表されるヒトの遺伝病には、遺伝子に含まれる3塩基の繰り返し（反復配列）が伸長する事によって起こる神経系の病気がある。このような病気は3塩基の反復配列が原因であるため「トリプレットリピート病」と呼ばれ、トリプレットリピートの長さがある一定の閾値を超える

と病気として発症するものが多い。トリプレットリピートの長さはヒトの集団中に多型として存在するし、病気を発症しない正常型のリピートはハンチントン舞踏病の場合、CAGリピートの単位の長さが9～30の範囲に収まるのに対して、病原型のリピートの場合40～121という長さになる。また病原型の場合、親から子に遺伝する際に急激な変異によってリピート長が大幅に伸びる事があり、病気

の発症を早める事が知られている。トリプレットリピートはコーディング領域においてアミノ酸リピートを形成する。形成されたアミノ酸リピートは、タンパク質の立体構造において不確定領域を形成しやすく、分子間相互作用を担う機能ドメインとしての役割がある事が示唆されている。

- (2) このようなアミノ酸リピートを持つ遺伝子を脊椎動物の各動物門(魚類、両生類、は虫類、鳥類、ほ乳類)の動物間で比較すると、より古い時代に出現した動物(魚類や両生類)はアミノ酸リピートを持つ遺伝子が少なく、アミノ酸リピートが短いものに対して、新しく出現した動物(ほ乳類)の方がアミノ酸リピートを持つ遺伝子が多く、リピート長が長い事が分かっている。この事に関して、ゲノム中のGCプレッシャーの上昇に伴って、リピート長が長くなるという仮説が出されている。この仮説では、ほ乳類は魚類や両生類に比べてゲノム中のGC含量が多いため、主にGCリッチなコドンで構成されるトリプレットリピートを多く持っていると説明される。
- (3) これまでのトリプレットリピート病に関する研究では、長いトリプレットリピートが細胞毒性を持つかどうか、どのくらい長いリピートを持てば病気が発症するのか、非コード領域のトリプレットリピート様の塩基配列はゲノム中にどれくらい存在するのか、といった研究が盛んに行われている。しかしながら、「なぜある一定の長さを持ったトリプレットリピートをヒトなどのほ乳類はゲノム中に維持しているか?」という点を明確に問題にして取り上げた研究はない。

2. 研究の目的

- (1) ヒトゲノム中(あるいはほ乳類動物のゲノム中)にこのようなトリプレットリピート(アミノ酸リピート)が維持されているのは、進化的な背景があるためと考え、本研究では、ヒトのトリプレットリピート病をダーウィン医学的な視点で捉えるために、ヒトまたはほ乳類におけるコーディング領域トリプレットリピートの進化的意義を明らかにする。すなわち、病気の原因になりうるにもかかわらずヒトまたはほ乳類が一定の長さのアミノ酸リピートをそのゲノム中に維持するようになった進化的なバックグラウンドや、そのメカニズムを明らかにする。

3. 研究の方法

- (1) ヒト、チンパンジー、アカゲザル、マウス、ラット等のほ乳動物について、全遺伝子のアミノ酸配列情報についてゲノムデータベースを検索し、トリプレットリピートを含む遺伝子を抽出した。また、オランウータン、テナガザル、ニホンザル、ヨザル、オオギャラゴについて、アミノ酸リピート部分の塩基配列をPCRダイレクトシーケンシングによって決定した。
- (2) ヒトのアミノ酸リピートを含む遺伝子の配列を問い合わせ配列として相同性検索をする事で、両生類の熱帯ツメガエルの全遺伝子配列からヒトのリピートを含む遺伝子に相同性の高い熱帯ツメガエルの遺伝子を抽出した。
- (3) ヒトの遺伝子を基に、タンパク質の立体構造の予測をし、それぞれのアミノ酸残基がどれくらい不確定領域を形成しやすいかを予測プログラムを用いて予測した。
- (4) アミノ酸リピートのうち、グルタミンリピートと特異的に結合するPQPB-1タンパク質と、トリプレットリピートを模したペプチドとの結合の強さを、表面プラズモン共鳴現象を用いて測定を行った。

4. 研究成果

- (1) 3-1 から得られたアミノ酸リピートの情報から、異なる種間でのアミノ酸リピートの長さの比較を行った。その結果、ほ乳類でリピート長が保存されているリピート(カテゴリCM)、霊長類ではリピート長が保存されているが、他のほ乳類ではリピート長が変化しているリピート(カテゴリCP)、霊長類においてもリピート長が変化しているリピート(カテゴリVP)に分類できた。
- (2) 3-3の結果からアミノ酸リピート中のアミノ酸残基のうち、不確定領域になりやすいアミノ酸残基の割合をDisorder Indexとして定義した。上記の3カテゴリ別のDIの分布は図1のようになった。DI=0とDI=1となるリピートを持つ遺伝子の割合をCM、CP、VPのカテゴリで比較をすると、顕著な違いがある事が分かった。カテゴリCPではDI=0のリピートを持つ遺伝子が相対的に少ないのに対して、カテゴリVPではDI=1のリピート

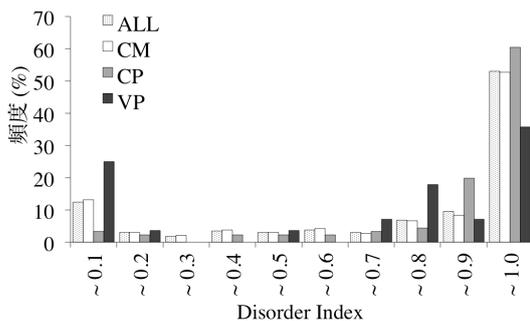


図1 アミノ酸リピートが持つ Disorder Index のカテゴリ別の分布

を持つ遺伝子が相対的に少なかった。非同義置換率と同義置換率の比 (Ka/Ks) をカテゴリ別に求めると、カテゴリ CP では、 $DI=1$ と $DI=0$ の遺伝子間に Ka/Ks の違いは無かったが、カテゴリ VP では $DI=0$ の遺伝子の方が $DI=0$ の遺伝子より Ka/Ks が大きく、進化速度が速い事が分かった。これまで、不確定領域になりやすい領域が進化しやすい事が言われていたが、霊長類の進化においてはそれとは逆に、不確定領域になりにくい領域の方が進化しやすい事が示された。

- (3) ヒトの遺伝病である脊髄小脳変性症の原因遺伝子の1つに Ataxin-1 がある。ヒトの場合、この遺伝子にアミノ酸リピートの1つである Q リピートが見られるが、このリピートが旧世界ザルに共通してみられるのに対し、新世界ザルでは単一アミノ酸リピートになっていない。そこで、3-4 の方法で、旧世界ザル型のリピート (QQQQQQQQQQ) と新世界ザル型のリピート (QQPQQPPPPQQ) を持ったペプチドを合成し、それぞれのリピートと PQBP-1 との結合能を実験的に測った。その結果、旧世界ザル型のリピートで見られた PQBP-1 タンパクとの結合能は新世界ザル型のリピートでは確認できなかった。図2で示すように旧世界ザル型のリピートを持つヒト遺伝子とそれを持たないマウスの遺伝子では Disorder Index に大きな違いがなく、タンパク質の構造が大きく変わらなくとも遺伝子の機能が変化しうる事が示された。
- (4) 3-2 の方法で得られた熱帯ツメガエルの遺伝子情報を基に、脊椎動物におけるアミノ酸リピートの進化の解析を試みた。ヒトのアミノ酸リピートを含む遺伝子とのホモログ遺伝子を熱帯ツメガエルゲノム中から抽出をしたところ、ヒト遺伝子の1/3についてホモログ遺伝子を抽出できた。これらの遺伝子に含まれるアミノ

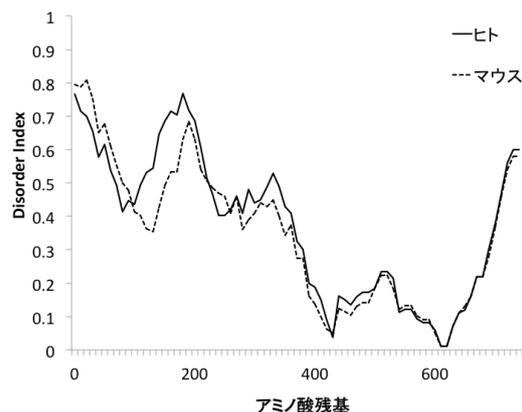


図2 Ataxin-1 遺伝子の Disorder Index

酸リピートを調べると、グルタミン (Q) が顕著に多い等の特徴はヒトのアミノ酸リピートを含む遺伝子と変わらなかった。一方で抽出できた遺伝子のうち、進化速度 (Ka/Ks) を評価できたのは半分程度で、残りの半分については塩基置換が起こりすぎていて評価ができなかった。アミノ酸リピートを含む遺伝子には転写調節因子等の機能的に重要な遺伝子が多く含まれるため、配列情報も保存されている事を期待していたが、進化速度を求められなかった遺伝子が多数あった事から、ほ乳類や霊長類におけるアミノ酸リピート長の比較と同様の事がヒトと両生類の比較では難しい事が分かった。

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)

- ① Kurosaki, T., Gojobori, J. & Ueda, S. "Comparative genetics of the poly-Q tract of ataxin-1 and its binding protein PQBP-1." *Biochem. Genet.* **50**, 309-317 (2012). 査読有り
DOI: 10.1007/s10528-011-9473-1
- ② Gojobori, J. & Ueda, S. "Elevated evolutionary rate in genes with homopolymeric amino acid repeats constituting nondisordered structure." *Mol Biol Evol* **28**, 543-550 (2011). 査読有り
DOI: 10.1093/molbev/msq225

〔学会発表〕（計 2 件）

- ① Gojobori J. and Ueda S. “Elevated Evolutionary Rate in Genes with Non-disordered Homopolymeric Amino Acid Repeats During Primate Evolution” Okazaki Biological Conference, 岡崎、2012 年 3 月
- ② Gojobori J. and Ueda S. “Evolutionary rates are elevated for genes with homopolymeric amino acid repeats constituting nondisordered structure” SMBE2011、京都、2011 年 7 月

6. 研究組織

(1) 研究代表者

五條堀 淳 (GOJOBORI JUN)

総合研究大学院大学・先導科学研究科・助教
研究者番号：00506800

(2) 研究分担者

なし

(3) 連携研究者

なし