

# ヒトゲノムに起こりうる有害な突然変異の予測

●今西 規<sup>1,2)</sup>

1) 国立遺伝学研究所 生命情報研究センター (旧所属)      2) 産業技術総合研究所 生物情報解析研究センター (現所属)

## ＜研究の目的と進め方＞

さまざまな遺伝性疾患の原因遺伝子の研究により、有害突然変異が生じるメカニズムには、単塩基置換、ハンチントン病に代表されるトリプレット（3塩基）リピートの伸長、染色体の構造異常などがあることが知られている。そこで本研究では、ヒトゲノム配列データを計算機上で解析して突然変異のホットスポットを探すことにより、ゲノム中に起こりやすい有害な突然変異を予測すること、さらに予測された有害突然変異を検出するための実験手法を開発・提案することを目的とする。

## ＜研究開始時の研究計画＞

ヒトゲノムのドラフト配列を用いて有害突然変異の予測を行う。次の手順で研究を進める。

- 1) 配列解析用計算機の設定とゲノム・mRNA配列データの整理
- 2) 起こりやすい突然変異の種類の推定
- 3) マウスの相同遺伝子との比較による高度保存遺伝子の抽出
- 4) 配列データ解析プログラムの開発
- 5) ヒトゲノムに起こりやすい有害突然変異の予測とそのデータベース作成
- 6) 有害突然変異の検出のための実験手法の開発と提案

## ＜研究期間の成果＞

ヒトゲノム配列や遺伝子配列を用いて、コーディング領域内での（1）C p G配列におけるCからTへの単塩基置換、（2）CAGなどのトリプレット・リピート、さらに（3）染色体脆弱部位にみられるCCGリピート、を網羅的に探した。

**結果（1）** C p GからT p Gへの塩基置換によってCGAコドンは終止コドンになりやすいと考えられるので、CGAコドンを多数持つ遺伝子を列挙し、突然変異による異常対立遺伝子が見つかるかどうかを調査した。

**結果（2）** ハンチントン病の原因遺伝子と同様なCAGリピートは他の多数の遺伝子にも見つかった（別表）。この中には、球脊髄性筋萎縮症、SCA1、歯状核赤核淡蒼球ルイ体萎縮症の原因遺伝子も含まれていた。この表に示した遺伝子は、突然変異によって有害な効果が現れる危険性がある。この他にも、まだ機能が明らかでない多数の遺伝子がCAGリピートを持つことを見いだした。これらの遺伝子の機能解析と突然変異対立遺伝子の探索により、新規の疾患原因遺伝子が見つかる可能性がある。

以上の結果を「突然変異ホットスポット・データベース」として公開するべく、準備を進めた。

## ＜達成できなかったこと、予想外の困難、その理由＞

所属研究機関の移籍により、研究を中断せざるを得ない状況になった。データベースも試験公開にとどまり、正式公開はされなかった。しかし、本研究のアイデアは、後に、マウスでのCAG/CAAリピートを持つ遺伝子の多様性に関する研究（論文1）や、ヒト完全長cDNAの統

合データベース構築の研究（論文2）などで生かされる結果となった。

## ＜今後の課題＞

今後は、予測された有害突然変異が実際にヒト集団中に見つかるかどうかの検証を行いたい。予測された有害突然変異をヒト集団サンプルに対して大量スクリーニングすることにより、もし有害突然変異が検出されれば、それは疾患原因や遺伝子機能の解明につながる可能性を秘めている。

## ＜研究期間の全成果公表リスト＞

- 1) Ogasawara M, Imanishi T, Moriwaki K, Gaudieri S, Tsuda H, Hashimoto H, Shiroishi T, Gojobori T, and Koide T (2005) Length variation of CAG/CAA triplet repeats in 50 genes among 16 inbred mouse strains. *Gene* 349:107-119.
- 2) Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, et al. (2004) Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones. *PLoS Biology* 2: 856-875.

表. CAGリピートを含む主なヒト遺伝子

UniGeneID	#repeats	Name	
Hs.79391	21	huntingtin (Huntington disease) (HD)	← ハンチントン病
Hs.239506	21	mab-21 cell fate-determining protein homolog (CAGR1)	
Hs.99915	20	androgen receptor (AR)	← 球脊髄性筋萎縮症
Hs.89230	19	calcium-activated potassium channel SK3 (KCNN3)	
Hs.1619	14	achaete scute homologous protein (ASH1)	
Hs.280413	13	nuclear receptor co-repressor 2 (NCOR2)	
Hs.198296	13	hbrm	
Hs.74520	12	SCA1 for ataxin	← SCA1
Hs.96253	11	alpha1A-voltage-dependent calcium channel	
Hs.233321	11	retina-derived POU-domain factor-1	
Hs.80961	10	DNA polymerase gamma, mitochondrial protein	
Hs.182280	10	myocyte-specific enhancer factor 2A (MEF2A)	
Hs.169488	10	DRPLA	← 歯状核赤核淡蒼球ルイ体萎縮症
Hs.108447	10	spinocerebellar ataxia 7	