

氏名(本籍)	いし かわ けん いち 石川 顕一 (千葉県)		
学位の種類	博 士 (医 学)		
学位記番号	博 乙 第 2066 号		
学位授与年月日	平成 16 年 10 月 31 日		
学位授与の要件	学位規則第 4 条第 2 項該当		
審査研究科	人間総合科学研究科		
学位論文題目	<b>Prediction of the Coding Sequences of Unidentified Human Genes. X. The Complete Sequences of 100 New cDNA Clones from Brain Which Can Code for Large Proteins <i>in vitro</i></b> (新規ヒト遺伝子のコーディング領域の予測。X. <i>in vitro</i> で大きな蛋白質をコードする正常ヒト脳由来新規遺伝子 100 個の全長塩基配列)		
主 査	筑波大学教授	理学博士	岡村直道
副 査	筑波大学教授	医学博士	加藤光保
副 査	筑波大学助教授	博士(医学)	櫻井武
副 査	筑波大学講師	博士(理学)	三輪佳宏

## 論文の内容の要旨

### [目的]

ヒトゲノムに書き込まれている情報は、mRNA を経て蛋白質となり、機能発現する。しかし、この蛋白質の情報をコードする領域は、全ゲノムの 3% 程度を占めるに過ぎない。それ故、ゲノムと生命現象をリンクして理解するためには、ゲノム構造を明らかにするだけでなく、より直接的に生体機能に関係している mRNA や蛋白質の一次構造情報を蓄積することが不可欠である。cDNA の全長塩基配列解析はまさにこのための研究である。

本研究遂行当時、既にいくつかの原核生物や酵母のゲノム全構造解析が行われていた。その結果から生体内に存在する全蛋白質のサイズ分布を予測すると、原核生物間では蛋白質サイズ分布が酷似しているにも関わらず、原核生物と酵母の間には蛋白質サイズ分布に大きな差が見られた。仮に 1,000 アミノ酸残基以上の蛋白質を大きな蛋白質と呼ぶと、原核生物では最も多い場合でも大きな蛋白質は全体の 3% 以下しか存在しなかった。それに対して酵母では、約 10% が大きな蛋白質に属していた。この結果より、大きな蛋白質には真核生物に特有のものが多くことが示唆された。

また、Protein Information Resource (PIR) データベース中に登録されている哺乳動物の大きな蛋白質の情報を機能分類した結果、形態形成や臓器機能に重要な働きをしていると思われるものがほとんどであり、高次の生体機能に関与する未知蛋白質の発見に長鎖 cDNA 解析が大きな貢献をする可能性を示唆していた。そこで、本研究では、高等生物特有の臓器であり、また高次機能の分子的理解が重要な課題となっている臓器として脳を選び、正常ヒト脳由来の長鎖 cDNA 解析を行った。

### [材料と方法]

大きな蛋白質をコードできる長鎖 cDNA クローンは、正常ヒト脳由来 cDNA ライブラリーをサイズ分画

し、そこからクローニングした。未知の長鎖 cDNA を得るための一次スクリーニングとして、まず *in vitro* のタンパク質合成を行い、合成される蛋白質のサイズを調べた。大きな蛋白質をコードしていた cDNA に対し、二次スクリーニングとして cDNA の末端配列を決定し、GenBank データベース (release 100.0) 検索を行い、既知遺伝子であるか否かの判断をした。その結果、「大きな蛋白質をコードし、かつ公的データベース上未知である」と判断した cDNA について、全長の塩基配列解析を行った。全長の塩基配列を決定した cDNA クローンは、組織ごとの発現解析と染色体マッピングを行った。蛋白質のコード領域上に複数個 open reading frame (ORF) の存在が予測された場合、ORF の確からしさを検討するために RT-PCR 産物を鋳型として直接塩基配列決定を行い、また GeneMark プログラムを用いて予測した ORF と比較した。

#### [結果と考察]

サイズ分画した正常ヒト脳由来 cDNA ライブラリー (平均鎖長 3.9kb, 4.5kb, 5.3kb, 6.1kb の 4 つのサイズ分画ライブラリー) からランダムに選んだ cDNA に対して *in vitro* のタンパク質合成と cDNA の末端配列決定を行い、大きな蛋白質をコードする未知遺伝子として 100 種類の cDNA について、全長の塩基配列解析を行った。

塩基配列決定後、複数個 ORF の存在が予測された cDNA クローンについては RT-PCR 産物の直接塩基配列決定を行った。その結果、4 つのクローン (KIAA0698, KIAA0700, KIAA0703, KIAA0710) はイントロンと思われる配列の挿入により ORF が分断されていると判断した。9 つのクローン (KIAA0699, KIAA0701, KIAA0702, KIAA0704 ~ KIAA0709) は何らかの人工的なエラーにより ORF が分断されていると判断した。最終的に配列決定した 100 種類の cDNA の平均鎖長は 4.9kb であり、平均 ORF 鎖長は 2.8kb (922 アミノ酸残基に相当) であった。個々の遺伝子の一次配列から ORF 鎖長、アミノ酸残基数、染色体マッピング情報を予測した。

DNA 配列、蛋白質配列、蛋白質モチーフの各データベースに対するコンピューター解析の結果、87 遺伝子について、既知遺伝子との相同性が認められた。これらを機能分類した結果、シグナル伝達/細胞間コミュニケーション関連遺伝子群に 28 遺伝子、細胞構造/モーター関連遺伝子群に 14 遺伝子、遺伝子発現制御に関連した遺伝子群に 12 遺伝子分類された。

また、今回新たに同定した 100 種類の新規遺伝子について、10 種類のヒト組織における発現プロファイルを明らかにした。

### 審査の結果の要旨

本研究は、ヒト脳で発現している比較的大きな分子量のタンパク質をコードしている 100 個の未知遺伝子 cDNA クローンについて、全長の塩基配列解析を行ったものである。審査対象論文の内容は、申請者が所属していた「財団法人かずさ DNA 研究所」で行われたプロジェクト研究 “Prediction of the Coding Sequences of Unidentified Human Genes” で得られた成果の一部をなすものであり、オーソドックスな方法によって得られたデータ目録的な性格が強い。しかしながら、ここで提示されている情報が今後の研究に多大な貢献をすることは明らかであり、また、このプロジェクト研究を通して中心的な役割を果たして多くの成果を上げた申請者の研究遂行能力と学識は高く評価できる。

よって、著者は博士 (医学) の学位を受けるに十分な資格を有するものと認める。