

IV. 成果報告

2. 新技術・新産業の創出に関する報告

マウス長鎖 cDNA の取得・構造解析とそのための効率化技術の開発

テーマ1の1 (J1)

①マウス長鎖 cDNA クローンの全塩基配列決定

(財) 千葉県産業振興センター (主任研究員) 古閑 比佐志

【目的と概要】

本サブテーマの目的は、他のサブテーマに先行し最も基本的なリソースたる cDNA を効率的に取得し、その構造解析をバイオインフォマティクスのノウハウを生かし達成することにある。具体的には、ヒト KIAA 遺伝子に相同性を有するマウス遺伝子(mKIAA)を取得するため、まず遺伝子ソースとなる各種 cDNA ライブラリーを作製した。引き続きライブラリーを展開して、個々の cDNA クローンの末端配列解析を網羅的に実施し、相同性検索の手法の一つである blast search を用いて mKIAA 遺伝子を同定した。地域分研究開発 (テーマ1-2) において当該 cDNA クローン選出後、プラスミド DNA を調整しショットガン法にて全長配列解析を行なった。選択したクローンのプラスミド DNA は全長配列解析に用いるのみならず、cDNA マイクロアレイ作製の際のソースとしても保存・管理した。全長配列解析結果は再度コンピュータ解析に供し、その相同性・全長性などを確認し最終的には公共データベースに登録した。プロジェクト半ばに mKIAA 遺伝子の取得が終了したことから、平成 17 年度には一旦このテーマを終了した。しかし、mKIAA 以外の長鎖 cDNA に研究の幅を広げることは、フェーズ III のける事業家に有利との判断から、平成 18 年度には同様の作業を再開し、最終的には 2679 クローンの全長配列解析を行なった。新規でないものに関しては ROUGE データベースでの公開は行なわなかったことから、ROUGE データベースにおいては 2169 クローンの公開となっている。

【研究成果の概要と今後の取り組み】

(1) mKIAA 遺伝子の取得及び全長配列解析

特に大きな蛋白質をコードする遺伝子の研究はその遺伝子取得の困難さ故立ち遅れていたが、かずさ DNA 研究所ではライブラリーの構築そのものに手を加えることでこの問題を克服し、1994 年から現在までに 2000 を越える新規ヒト長鎖 cDNA (> 4 kb) を解析してきた。これらの遺伝子にコードされる蛋白質は平均 949 アミノ酸残基で、1000 アミノ酸残基以上のヒト高分子蛋白質に限って言えば、かずさ DNA 研究所で見出されたものがその約 1/3 を占めていることになる。かずさ DNA 研究所で蓄積されてきたこれら遺伝子資源を如何に有効利用し蛋白質機能解析に反映させるかが、cDNA プロジェクト開始当時の目的でもあり我々に課せられた次なる使命である。このミッションに対して本事業では、そのマウス相同遺伝子をまず取得することで、モデル生物として最も有用な種に関してもリソースを整備することを第一の課題に挙げた。

方法としてはまず cDNA クローンに対して 5' または 3' 末端配列解析を網羅的に行い、事業開始前までにかずさ DNA 研究所で蓄積してきたものと合わせ約 170,000 のマウス cDNA クローン末端配列解析データのプールを構築した。この配列情報に対してヒト KIAA 遺伝子配列で BLAST サーチを行い BLAST スコア 120 以上のものをマウス KIAA 相同遺伝子候補として全長配列解析に供している。BLAST サーチで候補を選択できないものに関しては我々が hybridization 技術を利用して新たに開発したクローニング法[MUCH; Multiplex Cloning of Homologous Genes、小原令子ら、BioTechniques 36(5):798-804, 2004] や PCR 技術を用いて別途クローニングした (図 1)。これらの方法を用い現在までにヒト KIAA 遺伝子の約 97% に当たる 1984 マウス KIAA 相同遺伝子の全長配列解析を終えた。得られた配列情報は随時公共データベースに登録するとともに配列情報とそれから派生した様々な解析データを ROUGE(Rodent Unidentified Gene-Encoded)蛋白質データベース上でも公開している (<http://www.kazusa.or.jp/rouge/>)。我々はショットガン法を用いて全長配列解析を行っているが、その際配列解析に用いたショットガン断片はその後の抗体作製にそのまま用いられるように 0.8-1kb の断片としている。また大量サンプルを扱う際に陥りやすい人的ミスを最小限にすべく、ロボット化・作業フローのシステム管理も行った。一部の新規配列に関しては RT-PCR で組織特異的発現を解析後、特許申請も行った。

(2) mFLJ 遺伝子の取得及び全長配列解析

かずさ DNA 研究所では NEDO のプロジェクトの一環として、東京大学やヘリックス研究所とともに、KIAA 以外のヒト長鎖 cDNA の取得にも努めてきた。KIAA 遺伝子のマウス相同遺伝子がどうしても上記したような手法でも取得できない例があり（ヒト特異的な遺伝子も含む）、当初目標の 2000 に到達しない可能性も出てきたため、420 種ある FLJ 遺伝子のマウス相同遺伝子も取得することとした。その結果最終的には 185 種の mFLJ 遺伝子も取得し、目標値である 2000 を大きく上回る総計 2169 遺伝子を抗体作製の対象とすることが可能となった。

（3）mKIAA・mFLJ 類似遺伝子（ホモログ）の取得及び全長配列解析

将来的により多くの長鎖 cDNA のヒトとマウスのセットでの取得の優位性は明らかなため、さらに解析範囲を広げ mKIAA・mFLJ 類似遺伝子（ホモログ）の取得も行なった。これらは長鎖というだけでなく、構造的に類似した遺伝子群をまとめて収集するという観点からも極めて意義深い。たとえば類似したカイネースを収集すれば、特定の疾患に関わる分子や逆にその基質の同定に繋がる。最終的にはこれらの作業によって、さらに 254 種の遺伝子の取得に至った。これらのものの多くは既知遺伝子であったため、公共データベースや ROUGE データベースへの登録は控えた。

【主たる特許】（公開）

1. RasGEF モチーフを有する新規ポリペプチド及びそれをコードする DNA
特開 2004-081204
出願人：科学技術振興機構、かずさ DNA 研究所
2. BTB/POZ ドメインと Kelch 繰り返し配列を有する新規ポリペプチド及びそれをコードする DNA
特開 2004-073076
出願人：科学技術振興機構、かずさ DNA 研究所
3. HAT (Half-A-TPR) 繰り返しモチーフ及びプロリンに富む配列を有する新規ポリペプチド及びそれをコードする DNA
特許公開 2004-187668
出願人：科学技術振興機構、かずさ DNA 研究所
4. FHA、RING finger 及び D111/G-patch ドメインを有する新規ポリペプチド及びそれをコードする DNA
特許公開 2004-222680
出願人：科学技術振興機構、かずさ DNA 研究所
5. イムノグロブリン様繰り返しドメイン及びプロリンに富む配列を有する新規ポリペプチド及びそれをコードする DNA
特許公開 2004-283094
出願人：科学技術振興機構、かずさ DNA 研究所
6. FHA ドメイン及び膜アンカー領域を有する新規ポリペプチド及びそれをコードする DNA
特許公開 2004-337122
出願人：科学技術振興機構、かずさ DNA 研究所
7. UBA ドメインを有する新規ポリペプチド及びそれをコードする DNA
特許公開 2005-52003
出願人：科学技術振興機構、かずさ DNA 研究所

【2006 年までの主な成果】

- 1 Okazaki N, Kikuno R, Ohara R, Inamoto S, Hara Y, Nagase T, Ohara O, Koga H. Prediction of the coding sequences of mouse homologues of KIAA gene: I. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. 2002, 9(5):179-188
- 2 Okazaki N, Kikuno R, Ohara R, Inamoto S, Koseki H, Hiraoka S, Saga Y, Nagase T, Ohara O, Koga H. Prediction of the coding sequences of mouse homologues of KIAA gene: III. the complete nucleotide sequences of 500 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. 2003, 10(4):167-180

- 3 Okazaki N, Kikuno R, Ohara R, Inamoto S, Aizawa H, Yuasa S, Nakajima D, Nagase T, Ohara O, Koga H. Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. 2003, 10(1):35-48
- 4 Okazaki N, Kikuno R, Ohara R, Inamoto S, Koseki H, Hiraoka S, Saga Y, Seino S, Nishimura M, Kaisho T, Hoshino K, Kitamura H, Nagase T, Ohara O, Koga H. Prediction of the coding sequences of mouse homologues of KIAA gene: IV. The complete nucleotide sequences of 500 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. 2004, 11(3):205-218
- 5 Okazaki N, Kikuno R, Ohara R, Inamoto S, Koseki H, Hiraoka S, Saga Y, Kitamura H, Nakagawa T, Nagase T, Ohara O, Koga H. Prediction of the coding sequences of mouse homologues of FLJ genes: the complete nucleotide sequences of 110 mouse FLJ-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries. DNA Res. 2004, 11(2):127-135
- 6 Kikuno R, Nagase T, Nakayama M, Koga H, Okazaki N, Nakajima D, Ohara O. HUGE: a database for human KIAA proteins, a 2004 update integrating HUGEppi and ROUGE. Nucleic Acids Res. 2004, 32(Database issue):D502-504
- 7 Okazaki N, Imai K, Kikuno R, Misawa K, Kawai M, Inamoto S, Ohara R, Nagase T, Ohara O, Koga H. Influence of the 3'-UTR-length of mKIAA cDNAs and their Sequence Features to the mRNA Expression Level in the Brain, DNA Res., 2005, 12; 181-189

【図1】抗原蛋白質作製を見越した全長配列解析の流れ

