

# 論文の内容の要旨

論文題目        ヒト *CDC45L* 遺伝子の構造と転写調節

指導教官        青木 幸昌 助教授

東京大学大学院医学系研究科

平成9年4月入学

平成10年4月再入学

医学博士課程

生体物理医学専攻

学生証番号 87702

氏名            井垣 浩

ヒト *CDC45L* 蛋白質は DNA 複製起点に結合し、複製開始を制御している因子の一つである。DNA 複製は、誤動作が致命的となる細胞機能であり、従って *CDC45L* 遺伝子の発現は極めて厳密に調節されていると考えられる。有核細胞の遺伝子発現調節については、酵母やウイルスのプロモーターについて詳しく調べられているものの、ヒト遺伝子の転写調節について詳細に調べた報告は少ない。本研究では、*CDC45L* 遺伝子の転写調節に注目し、プロモーターと転写産物について詳細に検討した。

DNA 複製起点に集合し複製前複合体を形成する蛋白質の多くは酵母からヒトまで保存されている。そこで、出芽酵母 *Cdc45p* 蛋白質との相同性を手掛かりに、ヒト *CDC45L* 遺伝子由来 cDNA をクローニングした。まず *Cdc45p* 蛋白質のアミノ酸配列と相同性の高い cDNA を、National Center for Biotechnology Information の EST データベースに探し、互いに重複する配列を持つ複数の cDNA 断片を検出した。この cDNA の塩基配列をもとにプライマーを合成し、HeLa 細胞由来 cDNA ライブラリー中のヒトホモログ (*CDC45L* 遺伝子由来の cDNA) を PCR で増幅し、また、それと並行してヒト胎盤由来 cDNA ライブラ

リーからも通常のコロニーハイブリダイゼーション法により、*CDC45L* 遺伝子由来の全長 cDNA をクローニングした。また、この cDNA が位置するゲノムの塩基配列情報をもとにした DNA 断片をプローブにして、ヒト胎盤由来ゲノムライブラリーを探索し、*CDC45L* 遺伝子の転写調節領域を得た。

*CDC45L* 遺伝子は染色体 22q11.2 に位置しており、その 5'側には *UFD1L* 遺伝子の存在が示されていた。*UFD1L* 蛋白質はユビキチン付加後の蛋白質分解に関わることが知られている。クローニングした *CDC45L* の転写調節領域の塩基配列を明らかにしていくと、*CDC45L* と逆向きに、極めて近接して *UFD1L* 遺伝子が位置していることがわかった。この 2つの遺伝子の翻訳開始コドンは 884 塩基しか離れていなかった。*UFD1L* の翻訳開始コドンの 1 塩基上流を塩基番号 1 (nt 1) とし、*CDC45L* 方向に番号を付けると、*CDC45L* の翻訳開始コドン (ATG) は nt 885-887 に位置し、nt 1 から nt 884 の領域に *CDC45L* 及び *UFD1L* の発現を調節する主要なシス因子が存在することが推定された。

両遺伝子の転写開始点を 5'-RACE 法およびオリゴキャッピング法によって調べた。*UFD1L* 由来の mRNA は nt 69 (G)から転写されていた。*CDC45L* 由来 mRNA の大部分は nt 809 (G)から、一部は nt 382 (A)から転写されていた。従って、*UFD1L* は 1つの、*CDC45L* は 2つのコアプロモーター (下流側 : P<sub>CDC45L/major</sub> 及び上流側 : P<sub>CDC45L/minor</sub>) を持つことが示された。

一般に、転写開始点の約 40 塩基対上流から約 50 塩基対下流の領域はコアプロモーターと呼ばれる。この領域には基本転写因子群が結合することが知られており、*in vitro* ではコアプロモーターのみでも低レベルの転写が観察できる。コアプロモーターのすぐ上流には複数の転写因子が結合する数百塩基対の制御プロモーターと呼ばれる領域があり、基本転写因子群のコアプロモーターへの結合のしやすさを制御している。制御プロモーターとコアプロモーターが協同することにより、転写が能率よく起こる。そこで、nt 1 から nt 884 の DNA 断片をクローニングし、両端にホタルないしウミシイタケルシフェラーゼ遺伝子をつないだプラスミドを作成した。このプラスミドを HeLa 細胞に導入して一過性に発現するルシフェラーゼ活性を測定する方法で、コア及び制御プロモーターの領域を解析した。欠失変異を導入したプラスミドの転写活性を調べると、*UFD1L* 方向へは、nt 1-133 の DNA 断片で低レベルの転写があり、nt 1-487 の DNA 断片からは 884 塩基全体からとほぼ等しい転写が認められた。また、*CDC45L* 方向へは、nt 760-884 の DNA 断片から低レベルの、nt 355-884 の DNA 断片からは全長とほぼ等しい転写が見られた。上流側転写開始点から

の転写は全長 DNA 断片ではほとんど観察できないが、nt 488-884 の領域を欠失させると、高い転写活性がみられた。更に、nt 355-458 の短い断片から *CDC45L* 方向へ低レベルの転写が見られた。これらの成績は、3つのプロモーターとも、転写開始点に近い 104~133 塩基の領域内にコアプロモーターが存在することを示している。いずれのコアプロモーターも TATA ボックスを含んでいなかった。制御プロモーターは *P<sub>UFD1L</sub>* では nt 1-487、*P<sub>CDC45L/major</sub>* では nt 355-884 の範囲内に存在していた。*P<sub>CDC45L/minor</sub>* からの転写活性は 884 塩基対の DNA 断片ではほとんど観察できないが、*P<sub>CDC45L/major</sub>* とその制御プロモーターの大半を欠失させると高い活性が観察できるようになり、*P<sub>CDC45L/minor</sub>* の制御プロモーターと *P<sub>UFD1L</sub>* の制御プロモーターとは互いにほとんど重複していた。また、3つのプロモーターとも、コアプロモーターからの転写活性が制御プロモーターにより約 2 倍に増強されていた。

さらに、884 塩基長の DNA 断片に結合する可能性がある転写因子およびその結合配列を TFSEARCH プログラムで検索した。その転写因子結合配列の一部に塩基置換を導入することにより転写因子が結合できなくなる置換変異体を作成し、転写活性の差を調べた。転写に影響を与えている可能性が考えられた E2F、GATA-1、AML-1a (以上は *UFD1L* 方向)、E2F、CP2、Staf、CHOP-C/EBP  $\alpha$  2 量体 (以上は *CDC45L* 方向) の結合配列の変異では、明らかな転写活性の変化は見られなかった。しかし、*P<sub>CDC45L/minor</sub>* からの基底レベルの転写を反映する DNA 断片に対して、Staf もしくは CHOP-C/EBP  $\alpha$  2 量体の結合配列に変異を導入すると転写活性が消失した。従って、Staf および CHOP-C/EBP  $\alpha$  2 量体の転写因子は、*P<sub>CDC45L/major</sub>* からの転写には影響を与えていないが、*P<sub>CDC45L/minor</sub>* からの転写を制御している可能性が示唆された。

次に、cDNA ライブラリーからクローニングした *CDC45L* の cDNA を解析して、塩基配列の一部が異なる 512~610 アミノ酸の蛋白質をコードする 5 種類の cDNA を得た。これらの塩基配列を GenBank に登録されているヒトゲノムの塩基配列と比較したところ、これらはすべて 20 のエクソンで構成された単一の遺伝子から転写され、オルタナティブスプライシングによって作られた複数種類の mRNA に由来することがわかった。一部の cDNA は第 4、第 7、第 14 エクソンのいずれかを持たず、第 18 エクソンは cDNA により長さが異なっていた。即ち、*CDC45L* では 2 種類のプロモーターから、スプライシング部位の異なる多種類の mRNA が転写され、複数の蛋白質が作られることがわかった。

また、*CDC45L* から最も大量に作られる mRNA 由来の cDNA を *in vitro* 翻訳系に加

えて、蛋白質を合成した。この CDC45L 蛋白質は、大腸菌で作った GST-hMCM7 融合蛋白質および GST-p70 融合蛋白質と結合した。hMCM7 蛋白質は DNA 複製起点に結合するライセンス因子のひとつであり、p70 蛋白質は DNA ポリメラーゼ  $\alpha$  のサブユニットなので、MCM2-7 蛋白質複合体が複製開始転点に結合すると、CDC45L 蛋白質を介して DNA ポリメラーゼ  $\alpha$  が複製起点に誘導されると考えられる。

*UFD1L* 及び *CDC45L* の 3 つのコアプロモーターが、884 塩基対の極めて短い領域にあり、転写を制御する領域が互いに重なり合っていることは、これらの遺伝子の発現が共通の転写因子によって制御されている可能性を示している。また、*CDC45L* の 2 つのプロモーターは、転写される mRNA の 5'リーダー配列の違いを利用して、mRNA の半減期や翻訳効率を調節しているのかもしれない。このような遺伝子発現調節の機構を詳細に知ることは、遺伝子病の理解や遺伝子治療技術の開発に不可欠なものである。厳密な発現調節がなされていると考えられている TATA ボックスを持たないプロモーターの研究は遅れており、本研究の成果は今後の研究に有用な情報を提供している。

*UFD1L* と *CDC45L* が存在するゲノム領域は、DiGeorge 症候群患者で対立遺伝子の一方に高頻度に欠失が見られることから DiGeorge Critical Region (DGCR) と呼ばれている。この領域には多数の遺伝子が存在するため、原因遺伝子はまだ特定されていないが、*UFD1L* 及び *CDC45L* 蛋白質はどちらも個体の正常な発生と生存に必須であり、これらの発現異常が DiGeorge 症候群の原因である可能性を指摘する報告もある。遺伝子の比較的大きな欠失に注目するだけでなく、この 2 つの遺伝子の発現調節領域の小さな変異や修飾が遺伝子発現に大きく影響している可能性も考慮して、病因遺伝子の特定を進める必要があるだろう。