

ヒトおよびマウスLPC遺伝子のクローニング
と遺伝子構造解析

海老原 充

①

ヒトおよびマウス LPC 遺伝子のクローニング
と遺伝子構造解析

Imperial Cancer Research Fund
Department of Medical Oncology

海老原 充

1 目次

1 目次.....	1
2 緒言.....	3
3 ヒトLPC遺伝子のクローニングと解析.....	22
3-1 はじめに.....	22
3-2 材料および方法.....	23
3-2-1 ブレークポイントの同定.....	23
3-2-2 ブレークポイントのクローニング.....	24
3-2-3 ブレークポイントの解析.....	27
3-2-4 LPCの同定とクローニング.....	28
3-2-5 エクソン・イントロン構造の解析.....	32
3-3 結果.....	36
3-3-1 ブレークポイントの同定.....	36
3-3-2 ブレークポイントのクローニング.....	37
3-3-3 ブレークポイントの解析.....	38
3-3-4 LPCの同定とcDNAのクローニング.....	38
3-3-5 エクソン・イントロン構造の解析.....	50
4 マウスLPC遺伝子のクローニングと解析.....	54
4-1 はじめに.....	54
4-2 材料および方法.....	55
4-2-1 マウス染色体DNAのスクリーニング.....	55
4-2-2 欠失変異株の作成.....	59
4-2-3 DNAシーケンス.....	60
4-2-4 エクソン・イントロン構造の解析.....	61
4-2-5 マウスcDNAのスクリーニング.....	64
4-3 結果.....	70
4-3-1 マウス染色体DNAのスクリーニング.....	70
4-3-2 欠失変異株の作成.....	75
4-3-3 DNAシーケンス.....	77
4-3-4 エクソン・イントロン構造の解析.....	107
4-3-5 マウスcDNAのスクリーニング.....	109
5 考察.....	114

5-1	LPC 遺伝子とその構造	114
5-2	SUBTILISIN-LIKE SERINE PROTEASE	120
5-3	LPC と転座の関連	125
5-4	今後の研究	127
6	謝辞	128
7	参考文献	129

2 緒言

染色体 11 番長腕は、高い確率で欠失・転座が観察されることが知られている。主に細胞遺伝学的手法によって、さまざまな例が報告されている。特に、11q22-24 の領域は、乳癌、肺癌、直腸癌、悪性のメラノーマ、副神経節腫、白血病など多くの悪性腫瘍細胞で、欠失・転座が報告されている。

また、この領域には、血管拡張性失調症 (ataxia-telangiectasia) 関連遺伝子 (ATM)、白血病に関与する MLL がコードされていることも知られている。

以上のように、これらの染色体の欠失・転座は、白血病、リンパ腫、様々な癌と関連していることが示唆されている。遺伝子レベルでも、転座により新しい融合遺伝子ができ得ることが知られている。あるタンパク質の N 末端と別のタンパク質の C 末端が融合するのである。あるいは、ある遺伝子上流等に、別の遺伝子のプロモーター・エンハンサーなどの調節配列が転座してくる、またはプロモーターが欠失してしまうなどの影響により、その遺伝子の発現レベルが変化することもある。

融合タンパク質が生じる場合、通常、それぞれのイントロン内に、breakpoint が存在している。mRNA に転写されたのち、inframe にスプライシングされることにより、融合タンパク質が生じる。調節配列が転座してくる場合は、イムノグロブリンや T 細胞レセプターの rearrangement が関与している例が観察されている。染色体転座の研究によって、転座のパートナー遺伝子 (融合するそれぞれの遺伝子) には、かなりの特異性があることが明らかになっているが、ある特定の領域が関与する転座のパートナー遺伝子、あるいは領域は、一つであるとは限らない。すなわち、ある breakpoint は、複数かつある程度の特異性をもったパートナー遺伝子、あるいは領域を持っているのである。

このような breakpoint の一つに染色体 11q23 がある。11q23 は、様々な白血病細胞において転座が起こっていることが知られており、かつ 11q23 領域内の breakpoint も複数あることが知られている¹⁾。

t(4;11)(q21;q23)は、もともと最初に報告された 11q23 が関与する転座例であり、急性リンパ芽球白血病細胞 (acute lymphoblastic leukaemia, 以下 ALL) で見つかった^{2,3}。ALL 患者の子後は悪く、通常、非 B 細胞、非 T 細胞タイプであると共に、ターミナルデオキシヌクレオチダーゼ活性プラスである。t(4;11)(q21;q23)タイプの転座を持つ白血病は、ALL に分類されることが多く、かつイムノグロブリン遺伝子の rearrangement が関与していることが知られているが、例外も知られている。急性骨髄性白血病 (acute myeloid leukaemia, 以下 AML) がその例である。形態的に、これらの白血病細胞は、単球とリンパ球細胞からなっていることが示唆されている。また、t(4;11)転座を持つ急性白血病細胞を解析した結果から、rearrangement が起きたイムノグロブリン、およびリンパ球と骨髄球のマーカー遺伝子が発現していることが明らかになっている。

11q23 領域が関与している転座を解析するにつれ、急性白血病の約 5% のケースにおいて、11q23 転座が関与していることが明らかになってきた^{4,5}。また、11q23 が関連する転座には、少なくとも 25 ものパートナー領域が存在することが報告されている⁶。11q23 が関与する転座を生じる白血病には、ALL、AML、小児、遺伝的要因によるもの、そして他の癌に対する細胞毒性な治療法による二次的な白血病などが含まれる。さらに、11q23 が関連する転座は、非常に複雑であり、他の複数の転座が同時に起きていることもある。例えば、リンパ球白血病においては 4q21、1p32、19p13^{7,8,9,10,5}、骨髄球白血病では 1q21¹¹、2p21¹²、6q27¹³、9p22¹⁴、10p11¹⁵、17q25、19p13¹⁶などである。

HRX 遺伝子が同定されて以来、11q23 領域とそれに関与する転座の分子生物学的研究が飛躍的に進んだ。11q23 領域にマップされる HRX 遺伝子が成人、子供、二次的に問わず、ほぼすべての白血病に関与していることが明らかになったからである。そのパートナー遺伝子は、様々である¹⁷ (下表)。

表1 HRX 遺伝子が関与した転座とパートナー遺伝子

転座型	パートナー遺伝子	関連遺伝子
t(4;11)(q21;q23)	AF4	LAF4 (human)
t(9;11)(p22;q23)	AF9	Anc1 (yeast)
t(11;19)(q23;p13.3)	ENL	Anc1 (yeast)
t(1;11)(p32;q23)	epa15	eps15 (mouse)
t(6;11)(q28;q23)	AF6	cno (Drosophila)
t(10;11)(p12;q23)	AF10	cezf (C. Elegans), BR140 (human)
t(11;17)(q23;q21)	AF17	cezf (C. Elegans), BR140 (human)
t(11;19)(q23;p13.1)	ELL	p80 elongation factor
t(1;11)(q21;q23)	AF1q	?

<小児白血病>

さて、小児における 11q23 の異常と急性白血病関与は、1980 年代始めから知られていた。1980 年代の後半になって、さらに詳しい解析がなされると、白血病と 11q23 の異常（例えば、t(4;11)、t(9;11)、t(11;19)、t(11;17)など）に、強い相関関係があることが明らかになった^{15,18,19,20,21,22}。しかし、小児における 11q23 と白血病の強い相関に対して、子供（15 才ぐらいまで）の白血病においては、高い相関関係が示されていない（ALL で 5%、AML で 13% 程度である）。

小児、子供のどちらの場合も、11q23 の異常を伴う白血病の予後は思わしくない。芽球細胞においては、イムノグロブリンのH鎖に rearrangement が見られるとともに、骨髄球特異的なマーカー遺伝子の発現も見られる。また、AML や ALL など、すべての小児白血病の発症は、平均して 9 週から 16 週の間という報告もある。

現在、小児の AML、ALL における 11q23 の異常は、HRX 遺伝子が関与していることが明らかになっている^{23,24}。小児の ALL96 例を調べたところ、細胞遺伝学的に、50%に 11q23 の異常が観察された。しかし、サザンハイブリダイゼーションによる解析では、81%で HRX 遺伝子の rearrangement が観察された。

小児の白血病は、極めて短い潜伏期間（すなわち急性）を持つが、そのこと

は、子宮内ですでになんらかの異常が起こっていることを示唆していると考えられる。白血病の双子を詳しく調べたところ、どちらもまったく同じ HRX の rearrangement が起こっていることが明らかになっている。また、別の双子を調べた結果から、末梢血細胞において同じ t(11;19)(q23;p13.3)型の転座が起こっていることが明らかになった。このことは、小児の白血病における染色体異常は、子宮内ですでに起こっていることを示唆している。

さらに、双子でも異なる形の転座が起こっている例が知られている。この場合、HRX の点座が B 前駆体細胞か造血系の幹細胞でおこり、それが子宮内で他の胎児に移されたと考えられることができる。

<2 次的な白血病>

2 次的、あるいは治療によって生じた AML の増加が報告されている^{25,26}。主に別の異常に対し、化学治療を施した後に起きる可能性が高いとされている。骨髄異形成 (myelodysplasia) に関する治療を施した 91 例について調べたところ、共通して染色体 5 番と 7 番に異常が観察された。11q23 の異常も 9 例において観察されている。これら 11q23 の異常が生じた 2 次異常例では、骨髄異形成が再発する前に、極めて短い期間で白血病が発症している。12 例の小児 AML を調べた結果、5 例が 11q23 に異常があり、そのうち 4 例 ALL か non-Hodgkin's リンパ腫の細胞性塞栓に対する治療歴があった。様々な血液学的あるいは充実性腫瘍に対して epipodophyllotoxin etoposide や teniposide などの化学治療を施した成人・子供を含む 37 例について、治療後に生じた AML について調べた。その結果、これらの白血病は短い潜伏期間で起き、単球あるいは骨髄性単球が優勢で、かつ高頻度で 11q23 が関与した細胞遺伝学的な異常が観察された。充実性腫瘍に対して治療を行った 3365 例では、12 例の治療による骨髄性白血病が観察された。また、骨髄性細胞の染色体解析を行った 8 例では、2 例で完全な染色体 7 番の欠失、4 例では del(16)(q22)を伴う t(9;11)か t(8;21)、あるいは del(16)(q22)を伴わない t(9;11)か t(8;21)が見られた。epipodophyllotoxin の投薬を

受けた 2 例では 11q23 の異常が見られた。このデータや他に報告されているデータから考えると、epipodophyllotoxin 投薬を受けた患者は、アルキレート系の投薬と関与している 2 次白血病症候群とは一線を画する 2 次白血病を生じる危険性にさらされていることを示唆している。etoposide とアルキレート系の薬の作用機作は異なるようである。アルキレート系の薬は DNA と架橋を作る作用を介して働くのに対して、etoposide はトポイソメラーゼ II と相互作用することが知られている。様々な研究により、2 次白血病における 11q23 の異常は、de novo の白血病における 11q23 の異常と同様に HRX 遺伝子の rearrangement が関与していることが示唆されている。また、この事から、TRX 遺伝子内のトポイソメラーゼ II 結合部位が転座に関与しているかどうかという疑問が生じるが、この相関関係は不明である。

<HRX 遺伝子のクローニング>

さて、11q23 領域の転座の breakpoint を含む DNA 断片は、YAC クローンとして単離された²⁷。このクローンには、t(4:11)(q21;q23)、t(9:11)(p22;q23)、t(11:19)(q23;p13)で見られる breakpoint が含まれていることが、in situ ハイブリダイゼーションによって明らかにされた。また、この YAC クローンは、CD3 遺伝子を含むことが明らかになり、breakpoint が CD3 遺伝子近傍にあることが示唆された。同時に、breakpoint が CD3 遺伝子近傍に位置していることは、PFGE(Pulsed Field Gel Electrophoresis)によっても証明された。さらに、CD3 遺伝子から chromosome walk して得られた別の YAC クローンにも小児白血病細胞が持つ 11q23 の転座領域が含まれていることが明らかになり、転座に関与している領域が CD3 遺伝子近傍であることが再確認された。

さて、転座領域にマップされる遺伝子は、クローニングされる前から名前がつけられていて、MLL(mixed lineage leukaemia)、あるいは ALL1(acute lymphocytic leukaemia 1)などと呼ばれていた。そして、1992 年になって、ALL1²⁸、Htrx-1²⁹、TRX³⁰などの名前でも、3 カ所からクローニングが報告された。この論文では、HRX

遺伝子の名前で記述することにする。

HRX 遺伝子は、ゲノム上で約 100kb の大きさであり、21 のエクソンからなる。また、3969 アミノ酸残基、431kDa という極めて大きなタンパク質をコードしている。驚いたことに、転座の breakpoint は、100kb の領域のうち、わずか 8kb の BamH I 断片内に集中していることが明らかになった。この転座により、パートナー遺伝子との間に inframe の融合が起きていたのである。

HRX 遺伝子産物を解析すると、他の遺伝子産物との相同性が明らかになった。特に、*Drosophila melanogaster* の発達に關与する遺伝子 *trithorax(trx)* との相同性が高く、これが HRX 遺伝子の名前の由来でもある。もっとも相同性が高い領域は、アミノ酸レベルで 75% であり、この領域は *Drosophila* の Enhancer of Zeste にコードされているタンパク質の C 末端部分、position effect variegation (PEV) のサブプレッサー、および *Su (var)3-9* と *G9a* 遺伝子とも相同性が高い。この領域は、SET ドメインと呼ばれている (Suppressor of variegation, Enhancer of zeste, *Trithorax* から名付けられた)³¹。しかし、その機能は不明である。

SET ドメイン以外にも特筆すべきドメインがある。その一つが、HRX の中心近くに位置するシステインリッチ領域である。このシステインリッチ領域は、PHD あるいは、LAP ドメインと名付けられた Zinc finger ドメインであり、様々なタンパク質配列に見つかっている。この LAP は DNA メチルトランスフェラーゼと相同性があり、メチル化されていない DNA やメチル化された DNA を区別することに關与していると考えられている。転写の活性化ドメインは、この LAP ドメインの末端に位置している。また、転写のリプレッサードメインが LAP ドメインの真ん中に見つかっている。

さて、もう一つ特筆すべきドメインは、HRX 遺伝子産物の N 末端の AT フックである³²。この領域は、HMG-I(Y) の AT フックと相同性があるが、*Drosophila* の *trx* 遺伝子産物内には存在しない領域である。AT フックは、DNA の 2 重鎖の溝のうち、小さな溝(minor groove)に結合することが知られている。ただし、この結合によって、直接転写の活性化が行われるのではなく、何等かの因子の

活性を促進する働きがあると考えられている。

転座によって、リンパ系細胞の成熟等に必要とされるなんらかの機能を持つ TRX 遺伝子が、その機能を失うことになると考えられる。

<パートナー遺伝子のクローニング>

様々なグループにより、転座によって生じた融合遺伝子のパートナー遺伝子 (TRX 遺伝子と融合した遺伝子) のクローニングがなされてきた。

(1) t(4;11)(q21;q23)

t(4;11)(q21;q23)におけるパートナー遺伝子は、AF4 あるいは FEL と呼ばれている²⁸。AF4 遺伝子は、alternative splicing により 12kb と 10.5kb の 2 つの mRNA に転写され、1210 アミノ酸残基からなるタンパク質になる。また、AF4 は多くの組織で広く発現していることが知られている。

AF4 は nuclear targeting sequence (NTS) を持ち、セリン・プロリンリッチなタンパク質である。また、サブトラクション cDNA ライブラリーのスクリーニングから、LAF4 と呼ばれる AF4 と相同性が高い遺伝子がクローニングされた。LAF4 遺伝子産物も NTS 配列を持ち、核内に局在することが明らかになっている。

AF4、LAF4 遺伝子産物は、ともに GAL4 DNA 結合ドメインと相互作用した時に転写を活性化するドメインを持っていて、かつこの領域が HRX 遺伝子と融合した後も保持されている。この活性化ドメインが leukaemia の発症に関与していることが示唆されている。

(2) t(9;11)(p22;q23) と t(11;19)(q23;p13.1)

t(9;11)(p22;q23) と t(11;19)(q23;p13.1) 転座のパートナー遺伝子は、それぞれ AF9 と ENL 遺伝子である³⁰。AF9 と ENL は、アミノ酸レベルで 56% の相同性を持つ。t(9;11) と t(11;19) 転座は、AML と ALL の両方で見られる転座であり、leukaemia と深く関与していることが考えられる。

どちらの遺伝子産物も、NTS ドメインと C 末端に RNA ポリメラーゼ II に相

同的な配列を持っている。ENL は、C 末端の 90 アミノ酸残基領域に転写活性化活性を持ち、リンパ球と骨髄細胞の両者でその活性が観察されている。この転写活性化ドメインは、ENL、AF9 で保存されており、かつ HRX 遺伝子との融合遺伝子産物でも保存されている。このことにより、この転写活性化ドメインが、leukaemia の発症に関与していることが示唆される。

一方、出芽酵母からクローニングされた核タンパク質をコードする遺伝子 ANCl(actin non-complementing)が、AF9・ENL と高い相同性を保持していることが報告されている。この遺伝子の変異は、アクチン、細胞骨格などの構造に異常を示す。したがって、AF9 および ENL 遺伝子も、アクチンフィラメントの形成等に関与していることが考えられる。

(3) t(6;11)(q27;q23)

t(6;11)転座におけるパートナー遺伝子は、AF6 である³³。この遺伝子は、1612 アミノ酸残基からなり、アミノ酸レベルで *Drosophila* の *cno* 遺伝子産物と 40% の相同性があることが明らかになっている。*cno* 遺伝子産物は、Notch カスケードと他のシグナル伝達経路との相互作用を仲介すると考えられているタンパク質である。特筆すべきは、AF6、*cno* とともに、モータータンパク質キネシンプファミリーや、同じくモータータンパク質であるミオシンファミリーと相同性を持っていることである。この相同性を示す領域は、モータードメインではない。相同性をもったこの領域の機能は不明であるが、何等かのタンパク質との結合に関与していることが考えられている。したがって、AF6 遺伝子産物自身は、モータードメインを持っていないが、何等かの分子のトランスポートに関与していることも考えられている。

また、AF6、*cno* とともに DHR (GLGF) ドメインを持っており、このことは、細胞骨格と細胞膜の接点に局在する可能性を示唆している。

(4) t(1;11)(p32;q23)

t(1;11)転座におけるパートナー遺伝子は、AFIP である³⁴。この遺伝子は、896 アミノ酸残基からなるタンパク質をコードしている。AFIP 遺伝子産物は、ア

ミノ酸レベルで、マウスの eps15 遺伝子(epidermal growth factor receptor pathway substrate 15)と 88%の相同性を保持している。この eps15 遺伝子産物は、EGF レセプターによってリン酸化される細胞質内のターゲットである。eps15 は、N 末端にリン酸化ドメイン、その C 末側にフィラメントの形成に関与するミオシンのドメインと似たヘリックスドメインが存在している。さらに、DPF という 3 アミノ酸残基からなる繰り返し配列があり、この領域が機能ドメインであると考えられている。

(5) t(x;11)(q13;q23)

t(x;11)転座のパートナー遺伝子は、AFX1 遺伝子である³⁵。AFX1 遺伝子は、forkhead ファミリーに属する転写因子と高い相同性が見られる。しかし、forkhead ファミリー遺伝子と高い相同性を示す一方で、まったく相同性を持たない領域もあるので、特異なメンバーと考えられる。

(6) t(10;11)(p12;q23)と t(11;17)(q23;q21)

t(10;11)(p12;q23)と t(11;17)(q23;q21)転座におけるパートナー遺伝子は、それぞれ AF10 と AF17 である^{36,37}。AF10 と AF17 は、ともにシステインリッチな N 末端領域と C 末端のロイシンジッパー領域を持ち、高い相同性を示す。これらは、既に報告されている BR140 遺伝子とも高い相同性を示すが、BR140 遺伝子は、塩基性領域を持つ点が異なっている。また、C.elegans の CEZF 遺伝子も AF10 および AF17 と同様なシステインリッチ領域とロイシンジッパーを保持しており、これらは種を越えて高く保存されている遺伝子であると考えられる。

t(10;11)転座は、主に骨髄性単球ないし急性単球白血病細胞において観察される。t(10;11)転座を持つ 8 例の患者を調べたところ、AF10 のロイシンジッパー領域に HRX の N 末端が融合していることが明らかになっている。また、染色体 11 番の長腕の一部が、染色体 10 番に挿入されている異常も報告されているが、これも AF10 と HRX が関与していることが示唆されている。

リンパ腫細胞であり骨髄分化のモデル細胞としても使われる U937 細胞は、HRX と関与しない t(10;11)(p12;q14)転座を起こしている。この例では、染色体

10 番上の転座に関与している遺伝子は AF10 であるが、染色体 11 番上の遺伝子は、CALM と呼ばれる AP3 clathrin assembly protein であることが明らかになっている。これは、HRX 遺伝子以外の遺伝子と AF-10 が融合して最初の報告例である。

(7) t(11;19)(q23;p13.1)

t(11;19)(q23;p13.1)転座に関与している遺伝子は、ELL (MEN とも呼ばれる) である³⁵。この転座は、ENL 遺伝子に関与する t(11;19)(q23;p13.3)転座とは異なる。ELL 遺伝子は、塩基性が高いロイシンリッチなモチーフを持ち、これは pol(ADP-ribose) polymerase が持つ DNA 結合ドメインと相同性を示している。ELL 遺伝子は様々な組織で発現しているが、特に末梢血リンパ細胞や骨格筋、胎盤、精巣などでの発現レベルが高いことが知られている。近年、ELL は RNA polymerase II elongation factor であることが示されている。

(8) t(1;11)(q21;q23)

t(1;11)(q21;q23)転座に関与する遺伝子は、AF1q である³⁶。AF1q 遺伝子産物は、約 9kDa のタンパク質であるが、特に高い相同性を示す他のタンパク質が知られていない。また、胸腺細胞で強く発現しているが、末梢リンパ組織では発現していない。ただし、白血病細胞では発現していることが知られている。

(9) t(11;22)(q23;q11)

t(11;22)(q23;q11)転座に関与する遺伝子は、まだ同定されていないが、BCR 遺伝子よりセントロメアよりに breakpoint があることが明らかになっている。

(10) HRX の self fusion

AML において、HRX の self fusion が見られることが知られている³⁰。HRX 遺伝子が直鎖上に繰り返されており、2 から 6 回の繰り返し構造となっている。これらの繰り返し構造は、エクソン 6 のあとに別の HRX のエクソン 2 以下が融合する形になっており、かつ inframe な融合である。したがって、一部機能を持った HRX 遺伝子産物が発現しているのかもしれない。この self fusion には、Alu 配列が関与していると考えられている。

トリソミー 11 においては、2 本の正常な 11 番染色体に加え、HRX の self fusion を起こした 11 番染色体が存在する例が複数報告されており、トリソミー 11 と HRX の self fusion の関係が示唆されている。

<HRX 遺伝子の機能>

正常な HRX 遺伝子産物の機能は不明である。しかし、*Drosophila* と酵母における研究から、以下のことが示唆されている。

Drosophila の *trx* 遺伝子は、ホメオティック遺伝子の発現に必要であることが示されている。また、*brahma* (*brm*) 遺伝子は、酵母の転写活性化系に関与する SWI/SNF と高い相同性を持つ *trithorax* ファミリーに属する遺伝子である。この SWI/SNF 遺伝子産物は、クロマチン構造を修飾することにより、転写を促進すると考えられている。ヒトの *brm* 遺伝子である BRG1 は、ヒトの SWI/SNF 複合体と相互作用して、HRX とともに染色体構造を変化させるものと考えられる。HRX の N 末端に存在する AT フックは、HRX がクロマチンの構造変化に関与していることを示唆するものである。というのも、同じ AT フックを持つ HMGタンパク質は、ヒストン H1 と競合する形で染色体の構造変化に関与していることが明らかになっているからである。さらに、SWI/SNF 複合体は、ヒストンや HMG と相互作用し、転写を活性化することが明らかになっている。HRX 遺伝子を Knockout したマウス HRX(-/-) では、胎児期に致死となることがわかっており、このことは、HRX がホメオティック遺伝子の発現に関与していることを、間接的にサポートしていると考えられる。HRX(+/-) マウスでは、成長遅滞、骨格や造血系細胞の異常が観察されている。

HRX が関与した融合遺伝子の中には、クロマチンを修飾する因子を含んでいるものがある。BR140 は、TFIID の TAF250 サブユニットと高い相同性がありまた酵母の Anc1 は、TFIID と TFIIF のコンポーネントである。すなわち、HRX 融合遺伝子がクロマチン構造を修飾することによって、他の遺伝子の発現レベル・パターンを変化させることが考えられる。

HRX が他の多くのパートナー遺伝子と融合することが、白血病の発症に関与

しているのか、またこの変異が gain of function なのか loss of function なのかは、興味がある問題である。細胞遺伝学および分子生物学的研究によると、融合タンパク質は、常に N 末端側が HRX 遺伝子、C 末端側が他のパートナー遺伝子であることを明らかにした。このことは、重要なことは、HRX 遺伝子の C 末端が失われることであることを示唆している。すなわち、LAP/PHD ドメイン、trans-activation ドメイン、そして SET ドメインの欠失である。HRX 自身の self fusion が観察されることから、パートナー遺伝子が重要な働きをしていることは考えにくい。また、Knockout マウスの研究結果からも、転座によって loss of function が起こっていることが考えられる。

しかし、パートナー遺伝子が何等かの役割を持つという可能性もある。すべての融合タンパク質は、inframe であるからである。単に HRX 遺伝子の C 末端の欠失が重要なら、out of frame な融合遺伝子が見つかってもおかしくない。また、ENL と AF9、AF10 と AF17 のように、機能的に同じと考えられる遺伝子がパートナーになっていることも、パートナー遺伝子が何等かの働きをしていることを示唆するものである。さらに、転座と白血病細胞には、明らかな相関関係があることが示されている。HRX-AF10、HRX-AF6 は、AML-M4 と AML-M5 に見られるのに対して、HRX-AF9 と HRX-ENL は、リンパ、骨髄性、および未分化な白血病に見られるからである。さらに、AF10 は、U937 白血病細胞において、AF10 が他の遺伝子と融合していることが明らかになり、HRX が機能を失うことが重要ではないとも考えられる。

HRX-AF10 融合遺伝子に関する研究から、転座によって HRX 遺伝子にロイシンジッパーが融合していることがわかっており、このことは融合タンパク質がダイマーとなることが重要な意味を持つことを示唆している。多くの白血病細胞において、正常な HRX 遺伝子も発現していることを考えると、融合タンパク質とダイマーを形成するタンパク質は、正常な HRX 遺伝子産物なのかもしれない。

以上の知見からでは、HRX 遺伝子の変異が loss of function、gain of function

のいずれであるかの結論を導くことはできない。さらに解析が必要である。

<11q23 転座のメカニズム>

前述のように、11q23 にマップされる HRX 遺伝子が関与した転座の breakpoint は、極めて限られた領域 (8.3kb BamH I 断片) に集中している。この DNA 断片はすでに塩基配列が決定されており、イントロン 6 に 4 つの Alu 配列があるのを含めて、合計 8 つの Alu 配列が見つまっている⁴¹。10 例の詳細な breakpoint の解析から、5 例で breakpoint がイントロン 6 であることが明らかになった。このことは、Alu 配列が転座と関与していることを示唆する間接的な根拠であるが、一方でパートナー遺伝子の breakpoint 近傍に Alu 配列が見つからないこともわかっている。

t(4;11) と t(9;11) 転座において、転座に関与する両方の染色体の breakpoint 近傍にヘブタマー-ナノマーの組換えシグナルが見つまっている⁴²。このことは、V(D)J recombinase が転座に関与している可能性を示唆している。しかし、他の転座においてはヘブタマー-ナノマー配列が見つかっていないので、すべての転座がこれと同じメカニズムであるとは考えにくい。

また、別の t(4;11) 転座を解析した研究では、topoisomerase II 結合切断部位に見られる配列が見つまっている。2 次白血病を予防するために、topoisomerase 阻害剤が有効であることが知られている。すなわち、少なくとも 2 次的な白血病の発症に関しては、topoisomerase が関与していることが考えられる。HRX のイントロン 6 を含む 8.3kbp BamH I 断片の配列を調べたところ、topoisomerase II 結合切断部位と 1bp だけ違う配列が 11 箇所に見つまっている。

de novo の白血病と 2 次白血病では、同じ 8.3kbp の BamH I 断片内でも統計的に breakpoint が異なることが明らかになっている。de novo ではセントロメアよりの半分に集中しているが、2 次白血病ではテロメアよりの半分に集中している。

<Protein convertase>

翻訳後修飾 (post-translational modification) の概念は、今から 30 年前のインシュリンのプロセッシングの発見⁴⁴に端を発し、以来様々な研究がなされてきた。糖鎖付加、リン酸化、硫酸化などである。このようなタンパク質のプロセッシングの中で、もっとも基本的なものは、タンパク質の加水分解である。タンパク質の加水分解には、大きく分けて次の3つの分類がある。

- (1)粗面小胞体内でのシグナルペプチダーゼによるシグナルペプチドの除去。
- (2)細胞外に放出された後に加水分解。
- (3)細胞内での前駆体タンパク質 (proprotein あるいは precursor) のプロセッシング。

以後、この論文では、分類(3)に属するタンパク質の加水分解について述べることにする。

さて、多くの分泌タンパク質は、前駆体と呼ばれる大きな不活性型タンパク質として合成される。その後、特異的な部位が proprotein convertases (以下 PCs) により切断されて活性型になる。多くの場合の特異的な部位とは、2つの連続した塩基性アミノ酸 (Arg, Lys) の N 末端側である。

このようなプロセッシングを受けるタンパク質は、ペプチドホルモン・神経ペプチド・成長因子・膜結合型成長因子レセプター・血液凝固因子・細胞接着因子・ウイルスのエンベロープの糖タンパク質などであり、一般に調節タンパク質 (regulatory protein) と呼ばれている^{44,45}。

プロセッシングは、真核細胞の分泌過程で行われる。分泌タンパク質 (preproprotein) が粗面小胞体 (RER: rough endoplasmic reticulum) で合成されると同時にシグナルペプチドが切断され、生じた proprotein が、Trans Golgi Network (TGN) においてさらにプロセッシングを受け、成熟したタンパク質となる。また、プロセッシングの過程で生じる中間体 (pro 領域) の C 末端の 2 塩基性アミノ酸が切断され、カルボキシル基化 amidation されることが知られている。

さて、インシュリンタンパク質のプロセッシングの発見以後、このプロセッシングに関与する酵素に関する酵素についての報告は少なく、唯一酵母の Kex2

遺伝子にコードされる Kexin がプロホルモンのプロセッシングに関与していることが知られているにすぎなかった⁴⁶。

しかし、常に 2 塩基性アミノ酸が切断部位に存在していることから、何等かの酵素がこのプロセッシングに関与していることが示唆されてきた。また、切断される部位に構造的な多様性があることから、この過程に関与する酵素は複数であることも予想されてきた。切断部位に見られるアミノ酸配列は、Lys-Arg が約 47%、Arg-Arg が 22%、Lys-Lys・Arg-Lys が 3%、1 塩基性アミノ酸のみが 14%、4 つ以上の塩基性アミノ酸が 9%となっている。また、proprotein の 2 次構造がプロセッシングに重要であることも知られている。さらに、同じ遺伝子産物でも、組織特異的なプロセッシングを受けるものがあることも知られている。

1980 年代後半から、酵母の Kexin 様の酵素が哺乳類でも報告されるようになり、subtilisin 様セリンプロテアーゼ・ファミリーを形成していることが明らかになってきた。

現在までに報告されている哺乳類の subtilisin-like protease ファミリーは、furin、PC1/3、PC2、PC4、PC5/6、PACE4 (Paired basic Amino acid residue Cleaving Enzyme 4) の 6 種類である。

最初に報告された哺乳類の subtilisin 様セリンプロテアーゼは、furin である⁴⁷。furin をコードしている fur 遺伝子は、癌遺伝子 fes の上流 (fes upstream region) に見つかったことからその名前がつけられた。fur 遺伝子の発現パターンには組織特異性がなく、多くの細胞で発現していることが知られている。fur 遺伝子の発現レベルが上昇している細胞としては、primary non-small cell lung carcinomas、adenocarcinomas、squamous cell carcinomas⁴⁸が知られている。興味深いところでは、この fur 遺伝子は、複数の転写開始点を持っていることである。また、組織特異的な alternative splicing も観察されている。

furin が認識して切断する配列は、Arg-x-Lys/Arg-Arg であり、pro-TGF β 1⁴⁹、HIV の gp160⁵⁰、pro-NGF などが基質候補として知られている⁴⁵。しかし、どの

基質も *in vitro* 系で調べられたものであり、*in vivo* での基質となりうるかは明らかになっていない。近年、HIV の GP160 の例では、*furin* よりもより効率的にプロセッシングする別の酵素の存在が示唆された⁵¹。

furin 遺伝子をはじめとする *subtilisin-like proteases* は、哺乳類のみならず、広く種を越えて保存されていることが知られている。これまで、*anglerfish*、*Xenopus laevis*、*Drosophila melanogaster*、*Caenorhabditis elegans*、*Lymnea Astagnalis*、*Hydra vulgaris* などで報告されている。

これらの *subtilisin-like protease* の基本構造は次のようになっている。N 末端から、pre 領域 (シグナルペプチド)、pro 領域、catalytic domain、middle domain (P domain)、そしてそれぞれの *subtilisin-like protease* によって異なる C 末端領域である。pre 領域は、分泌タンパク質などに共通に見られる配列であり、ER 内で、シグナルペプチダーゼにより切断される (pre 領域が切断されたタンパク質を proprotein と呼ぶ)。その後、proprotein 内の pro 領域が切断されて、成熟した *subtilisin-like protease* となる。

subtilisin-like protease の catalytic domain は、種間あるいは *subtilisin-like protease* 間で極めてよく保存されていることが知られている。特に、セリン、ヒスチジン、アスパラギン酸などの 4 アミノ酸残基およびその周辺領域は相同性が高い⁵²。ただし、ただし、PC2 においては、アスパラギン酸がアスパラギンに置き換えられている。また、この catalytic domain は、*subtilisin-like protease* 間でもっとも保存されており、酵母の *kex* とヒトの *furin* 間でも 47% のアミノ酸残基が保存されている。哺乳類で最初に見つかった *furin* と酵母 *kex* 遺伝子で 47% のアミノ酸が保存されていることから、この catalytic domain、特によく保存されているセリン、ヒスチジン、アスパラギン酸周辺の塩基配列を用い、degenerate primers を設計して新しい *subtilisin-like protease* をクローニングする試みがなされ、PC1⁵³、PC2^{51,54}、さらに PACE4⁵⁵、PC4^{56,57}、PC5/PC6^{58,59} がクローニングされた。

middle domain は、酵母 *kex* の P ドメインであり、前述の catalytic domain と比

較すると相同性が低く、酵母とヒトでは31%のアミノ酸残基が保存されている。P domain は、kex においては、酵素活性を保持するために必要であることが知られている。

middle domain より C 末端側の領域は、subtilisin-like protease によってかなり異なり、ほとんど相同性が見られない。ただ、furin、PACE4、PC5/PC6 においては、システインリッチな構造が観察され、また、furin と PC5/PC6 は、transmembrane domain を持っていることが明らかになっている。

以下に哺乳類 subtilisin-like protease の基本構造を記す。

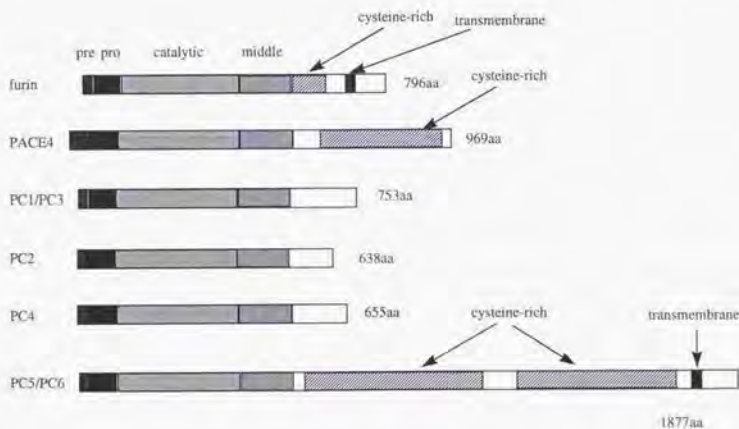


図1 subtilisin-like protease ファミリーと構造

これまでに報告されている哺乳類の subtilisin-like protease の発現パターンを調べると、次のようなことがわかる。

PC1/3 と PC2 は brain 特異的な発現を示すが、PACE4、furin は、調べられたすべての細胞で発現している。また、PC4 は、testis 特異的である。PC5/PC6 は、

intestine, brain, testis, ovary などでは発現しているが, pancreas や kidney では発現が見られない。

哺乳類の subtilisin-like protease の発現パターンを表にまとめる。

表2 subtilisin-like protease の発現

	PC1/3	PC2	PC4	PC5/6A	PC5/6B	Furin	PACE4
Pituitary	+	+					
Hypothalamus	+	++		+			
Anterior lobe	++	-		+			
Intermediate lobe	+	++		-			
Posterior lobe	-	(-)					
Adrenals	+	(-)		++			
Brain	(-)	++	-	+	-	++	++
Heart	-	-	-	-		++	++
Pancreas	-	-	-	-		+	+
Kidney	-	-	-			++	+
Liver	-	-	-	(+)		++	++
Muscle	-	-					+
Lung			-	+		++	+
Spleen	-	-		+			
Islet	+	+					
Intestine	-	-	-	+	+	++	
Testis			++	(+)		++	
Ovary			-	+			
Placenta							++

subtilisin-like protease の細胞内局在を調べた結果によると, これらは TGN (trans-Golgi network) に局在することが示唆されている⁶⁶。furin タンパク質の C 末端には, TGN targeting signal の存在が示唆されている。YKGL と CPSDSEDEG である。これらのシグナルを欠失させると, 免疫組織学的に検出される furin タンパク質の局在が変わることが報告されている。

また、furin 遺伝子をノックアウトしたマウスは致死になることがわかっており、発生の段階でも重要な働きをしていることが示唆されている (personal communication)。ただし、上述のように、in vivo での基質が何であるか、またそれぞれの subtilisin-like protease において異なった基質特異性を持っているのかは定かではない。合成ペプチドや HIV の GP160 タンパク質を用いた研究では、subtilisin-like protease によって、基質を切断する活性に差が見られることが報告されているが、in vivo でどのようにになっているかは定かではない。

この論文では、non-Hodgkin's リンパ腫にまれに起こっている t(11;14)転座部位にマップされる遺伝子のクローニング、およびその解析を行ったので報告するここで報告する遺伝子は、上記の subtilisin-like protease ファミリーに属する遺伝子 LPC (Lymphoma Proprotein Convertase) であり、哺乳類で 7 番目の subtilisin-like protease ファミリーに属する酵素である。

3 ヒトLPC遺伝子のクローニングと解析

3-1 はじめに

11q23 転座は、多くの白血病細胞において観察される転座である。特に、ほぼすべての細胞において、TRX 遺伝子が転座に関与していることがわかっている。しかし、リンパ腫細胞においては、11q23 転座が関与する例はあまり知れていない。この章では、non-Hodgkin's リンパ腫細胞のカリオタイプ解析の結果から得られた t(11;14)型の転座に関与する遺伝子の同定、およびクローニングを報告する。

3-2 材料および方法

3-2-1 ブレークポイントの同定

non-Hodgkin's 型リンパ腫細胞のカリオタイプの解析結果から、このリンパ腫患者由来の細胞には、t(11;14)(q23;q32)転座が起こっていることが判明していた⁶⁾。14q32 には、イムノグロブリン H 鎖がコードされており、イムノグロブリンの組換えが転座に関与していることが考えられたため、H 鎖の各部位を用いてサザンハイブリダイゼーションを行った。

サザンハイブリダイゼーションは、以下のように行った。

ヒト高分子量 DNA 10 μ g を 50units の BamH I, EcoR I, Hind III, Pst I, Bgl II で消化し、0.7%のアガロースゲル (25V, 一晚) で、電気泳動した。その後、エチジウムブロマイド 0.5 μ g/ μ l を含む 1 \times TBE バッファーに 15 分間浸して染色し、泳動を確認するとともに、スケールをあてて写真を撮った。0.25N HCl 溶液に 30 分間、0.4N NaOH 溶液に 20 分間浸した後、中和することなく 0.4N NaOH 溶液を用いて、アルカリトランスファーを行った。トランスファーを 2 時間行った後、2 \times SSC 溶液でメンブレンをリンスし風乾した。メンブレンは、Hybond N+ (Amersham) を用いた。

プローブの標識は、100 $^{\circ}$ C で 10 分間加熱して変性させた後、ランダムプライマーラベリングキット (Amersham) を用いて行った。標識には、約 50ng の DNA を用いた。標識後、スパンカラムにより標識された DNA プローブを精製した。

メンブレンを以下の表に示したハイブリダイゼーション溶液中で 65 $^{\circ}$ C、1 時間プレハイブリダイゼーションした後、100 $^{\circ}$ C、5 分間加熱後氷冷した DNA プローブを加えて、65 $^{\circ}$ C で一晚ハイブリダイゼーションを行った。

表3 ハイブリダイゼーション溶液の組成

10×Denhart's
4×SSC
1% SDS
100 μg/ml salmon sperm DNA

ハイブリダイゼーション後、次の条件でメンブレンを洗浄した。

2×SSC/0.1% SDS 室温、5分

2×SSC/0.1% SDS 室温、5分

1×SSC/0.1% SDS 65°C、10分

その後、フィルムカセット内で-70°C、一晚露光させた。

3-2-2 ブレークポイントのクローニング

ブレークポイントのクローニングは、以下のように行った。

まず、t(11;14)(q23;q32)転座を持つリンパ腫細胞（ICRF 細胞サービス部により樹立）、および末梢血から分離した細胞を用い、以下のようにコスミドライブラリーを作製した。

10⁹細胞を遠心で試験管に集め、氷冷した PBS で2回洗った。その後、10ml の消化バッファー（100 mM NaCl、10 mM Tris-Cl, pH 8、25 mM EDTA, pH 8、0.5% SDS、0.1 mg/ml proteinase K）を加え、50°Cで一晩インキュベートした。その後、等量のフェノール：クロロホルム：イソamilアルコール（24:24:2）で3回抽出した。1/2量の7.5M 酢酸アンモニウムと2倍量のエタノールを加えて2,000gで5分間遠心した後、上清を捨て70%エタノールで洗浄した。その後、1mlのH₂Oに溶解させた(1mg/ml)。

コスミドライブラリー用のDNAを得るために、ライブラリーの材料となる染

染色体DNAの部分消化の条件設定を行った。用いたコスミドは、lorist4である。

10 μ gのDNAをエタノール沈殿させた後、90 μ lのH₂Oと10 μ lの10 \times 制限酵素バッファーを加えた。5本の微量遠心管に1~5までの番号をつけ、1の微量遠心管に30 μ l、2~4には20 μ l、5には10 μ lのDNA溶液を分注した。まず、1の微量遠心管に15unitsの制限酵素Sau3A Iを加えた。1の微量遠心管から10 μ lを2へ移し、2の微量試験管から10 μ lを3の微量遠心管へ移すことを5の微量遠心管まで繰り返した。表を参照。

表 4 染色体 DNA の部分消化

微量遠心管の番号	1	2	3	4	5
DNA量(μ g)	2	2	2	2	2
制限酵素量(units)	10	3.3	1.1	0.4	0.2
濃度(units/ μ g DNA)	5	1.7	0.6	0.2	0.1

37 $^{\circ}$ Cで15分間インキュベートした後、0.5% アガロースを用いたゲル電気泳動を使用して適当な反応条件を決めた。反応条件が決定後、1mgのDNAを用いて制限酵素処理を行った。反応溶液は10mlで行った。反応開始後、13分で3.3ml、15分で3.3ml、17分で3.3mlを取り、反応を停止した。フェノール/クロロホルムで抽出後、エタノール沈殿により回収した。

回収したDNAを用いて、さらにシュークロース勾配によるサイズセレクションを行った。

10~40%のシュークロース勾配を超遠心チューブ内に作製し、これに65 $^{\circ}$ Cで5分間加熱したDNAを重層した。その後、150,000gで24時間、20 $^{\circ}$ Cで超遠心した。ペリスタポンプを使って、チューブの底から750 μ lずつ微量遠心管に分注し、各サンプルの中から40 μ lずつを使って、0.5%アガロースゲル電気泳動により、適当な大きさのDNAを含むものを選択した。

その後、以下の表のようにライゲーション反応を行った。

表5 ライゲーション反応

	1	2	3	4	5
ベクター(100ng/ μ l)	1	1	1	1	1
挿入 DNA (20ng/ μ l)	2				
挿入 DNA (10ng/ μ l)		2	1	0.5	
挿入 DNA (1ng/ μ l)					1
10×ライゲーションバッファー	0.5	0.5	0.5	0.5	0.5
H ₂ O	1	1	2	2.5	2
200units/ μ l T4 ligase	0.5	0.5	0.5	0.5	0.5

10×ライゲーションバッファー ; 660mM Tris·HCl (pH7.6) /10mM スペルミジン/100mM MgCl₂/150mM DTT/2mg/ml BSA(DNase/RNaseフリー)/5mM ATP

ライゲーション後、Gigapack gold(Stratagene)を用いてパッケージングした。その後、E.coli に transfect した。

245×245mm のプレートに約 1×10^5 コロニーをプレーティングし、常法により、Hybond N+へトランスファーし、アルカリで固定した。アルカリ固定は次のような行った。

メンブレンを 3MM ペーパーの上に置き、最低 10 分間乾燥させた。その後、

0.2N NaOH/1.5M NaCl 1 分間

1.5M Tris·HCl/2×SSC 1 分間

2×SSC 1 分間

処理し、3MM ペーパー上で乾燥させた。

その後、をプローブとして、以下のようにコロニーハイブリダイゼーションを行った。

プローブの標識は、ランダムプライマーラベリングキット (Amersham) を用いて行った。標識には、約 50ng の DNA を用いた。標識後、スパンカラムにより標識された DNA プローブを精製した。

メンブレンを以下の表に示したハイブリダイゼーション溶液中で 42°C、1 時間プレハイブリダイゼーションした後、100°C、5 分間加熱後氷冷した DNA プローブを加えて、42°Cで一晩ハイブリダイゼーションを行った。

表6 ハイブリダイゼーション溶液の組成

50% Formamide
5×Denhart's
5×SSC
0.1% SDS
100 μg/ml salmon sperm DNA

ハイブリダイゼーション後、次の条件でメンブレンを洗浄した。

2×SSC/0.1% SDS 室温、5分

2×SSC/0.1% SDS 室温、5分

0.2×SSC/0.1% SDS 42°C、15分

0.2×SSC/0.1% SDS 42°C、15分

その後、フィルムカセット内で-70°C、一晩露光させた。

3-2-3 ブレークポイントの解析

得られたコスミドクローンを用い、部分的な制限酵素マップを作製し breakpoint を含む DNA 断片、および正常な DNA 断片をサブラスミド pBluescript SK(-)にサブクローニングした。その後、これらの DNA 断片の DNA シーケンスを行った。

DNA シーケンスは、ABI373A を使い、ダイターミネーター法により行った。ダイターミネーター法によるシーケンス反応は、DNA Thermal Cycler Model 480

(Perkin-Elmer)を用いて以下の条件で行った。

表7 シーケンス反応溶液

鋳型 DNA	8 μ l (0.2~0.5 μ g)
ターミネーター反応 mix	8 μ l
プライマー	4 μ l

反応サイクル

96°C 5分 1 サイクル

↓

96°C 30秒

50°C 15秒 25 サイクル

72°C 4分

3-2-4 LPC の同定とクローニング

COS6 クローンから単離した BamH I/Bgl II 断片を、エクソントラッピング用ベクター pSP13 ヘサブクローニングした後、Buckler の方法によりエクソントラッピングを行った⁹。以下に、エクソントラッピングの概要を示す。ラビット β グロビンの 5' および 3' スプライシング部位の間にエクソンが含まれていると思われる染色体 DNA を挿入し、これを CHO 細胞にトランスフェクトする。その後、細胞質 mRNA を調製し、これを用いて cDNA を合成する。この cDNA には、CHO 細胞由来の遺伝子が大部分であるので、ラビットの β グロビン遺伝子領域にアニーリングするプライマーセットを作製して、PCR 反応を行う。その後、TA-cloning kit (Invitrogen) により、プラスミドへクローニングを行った。

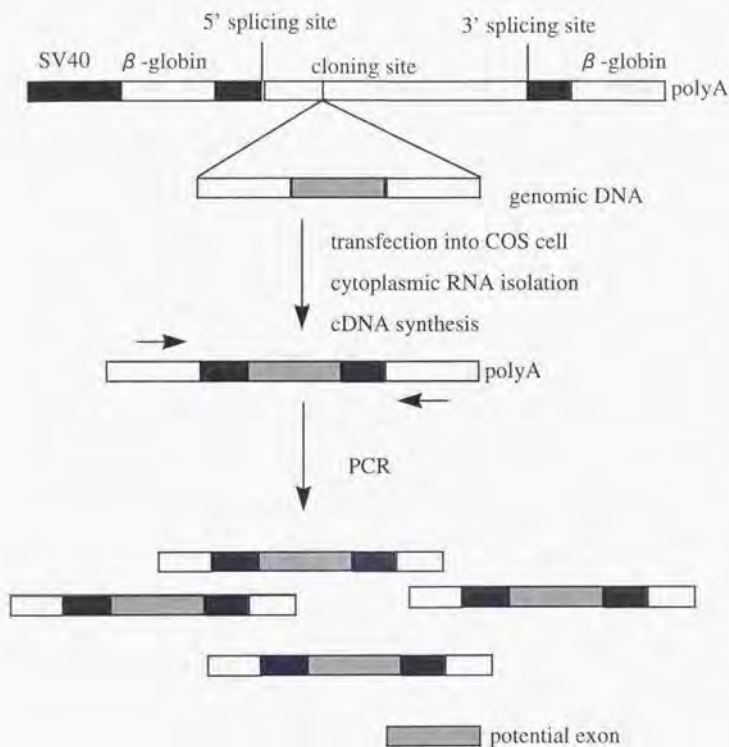


図2 エクソントラッピングの概略

得られたクローン *etp12* を用い、以下のように、Jurkat 細胞から作製した cDNA ライブラリーをスクリーニングした。このライブラリーは、J. Dunne 博士から供与された。

245×245mm のプレートに約 1×10^5 コロニーをプレーティングし、常法により、Hybond N+へトランスファーし、アルカリで固定した。アルカリ固定は次のように行った。また、メンブレンは、8 枚調製した。

メンブレンを 3MM ペーパーの上に置き、最低 10 分間乾燥させた。その後、

0.2N NaOH/1.5M NaCl 1分間

1.5M Tris·HCl/2×SSC 1分間

2×SSC 1分間

処理し、3MM ペーパー上で乾燥させた。

その後、*ctp12* をプローブとして、以下のようにコロニーハイブリダイゼーションを行った。

プローブの標識は、ランダムプライマーラベリングキット (Amersham) を用いて行った。標識には、約 50ng の DNA を用いた。標識後、スパンカラムにより標識された DNA プローブを精製した。

メンブレンを以下の表に示したハイブリダイゼーション溶液中で 42°C、1 時間プレハイブリダイゼーションした後、100°C、5 分間加熱後氷冷した DNA プローブを加えて、42°Cで一晩ハイブリダイゼーションを行った。

表8 ハイブリダイゼーション溶液の組成

50% Formamide
5×Denhart's
5×SSC
0.1% SDS
100 μ g/ml salmon sperm DNA

ハイブリダイゼーション後、次の条件でメンブレンを洗浄した。

2×SSC/0.1% SDS 室温、5分

2×SSC/0.1% SDS 室温、5分

0.2×SSC/0.1% SDS 42°C、15分

0.2×SSC/0.1% SDS 42°C、15分

その後、フィルムカセット内で-70°C、一晩露光させた。

単離されたクローンをを用いて、DNA シーケンスを行った。

DNA シーケンスは、ABI373 を使い、ダイターミネーター法により行った。
ダイターミネーター法によるシーケンス反応は、DNA Thermal Cycler Model 480 (Perkin-Elmer)を用いて以下の条件で行った。

表9 シーケンス反応溶液

鋳型 DNA	8 μ l (0.2~0.5 μ g)
ターミネーター反応 mix	8 μ l
プライマー	4 μ l

反応サイクル

96°C 5分 1 サイクル

↓

96°C 30秒

50°C 15秒 25 サイクル

72°C 4分

得られた cDNA を用いて、常法によりノーザンハイブリダイゼーションを行った。用いた mRNA は、ヒト spleen, thymus, prostate, testis, ovary, small intestine, colon, peripheral blood leukocyte, マウス heart, brain, spleen, lung, liver, skeletal muscle, kidney, testis である。

その後、得られた cDNA をプローブとして、次のように PI ライブラリーをスクリーニングした。PI ライブラリーは、ICRF の DNA ライブラリーサービス部から供与されたメンブレンフィルターを用いた。

プローブの標識は、ランダムプライマーラベリングキット (Amersham) を用いて行った。標識には、約 50ng の DNA を用いた。標識後、スパンカラムにより標識された DNA プローブを精製した。

メンブレンを以下の表に示したハイブリダイゼーション溶液中で 42°C、1 時間プレハイブリダイゼーションした後、100°C、5 分間加熱後氷冷した DNA プローブを加えて、42°Cで一晩ハイブリダイゼーションを行った。

表10 ハイブリダイゼーション溶液の組成

50% Formamide
5× Denhart's
5× SSC
0.1% SDS
100 μg/ml salmon sperm DNA

ハイブリダイゼーション後、次の条件でメンブレンを洗浄した。

2×SSC/0.1% SDS 室温、5分

2×SSC/0.1% SDS 室温、5分

0.2×SSC/0.1% SDS 42°C、15分

0.2×SSC/0.1% SDS 42°C、15分

その後、フィルムカセット内で-70°C、一晩露光させた。

得られた P1 クローンは、30~50ml の LB 培地を用いて培養し、アルカリ SDS 法により DNA を精製した。

3-2-5 エクソン・イントロン構造の解析

得られたヒト LPC cDNA の塩基配列から、PCR プライマーを作製し、以下の条件により PCR を行った。テンプレートは、先にクローニングされた LPC を含む P1 クローンを用いた。

表 11 反応溶液組成

鋳型 DNA (ヒト LPCPI クローン)	10.0 μ l
10 \times PCR バッファー	5.0 μ l
25mM MgCl ₂	5.0 μ l
2mM dNTPs (each)	5.0 μ l
プライマー 1 (20mM)	2.5 μ l
プライマー 2 (20mM)	2.5 μ l
H ₂ O	19.0 μ l
1u/ μ l Taq DNA ポリメラーゼ	1.0 μ l

反応サイクル

94°C 5分

60°C 2分 1 サイクル

72°C 10分

↓

94°C 1分

60°C 1分 30 サイクル

72°C 1分

↓

72°C 10分 1 サイクル

PCR 終了後、0.7 および 1.2%アガロースゲル電気泳動で増幅産物を解析するとともに、増幅産物をアガロースゲルから切り出し、GeneClean II kit (Bio101)により精製した。

精製した各 DNA 断片は、TA Cloning kit (Invitrogen)により pCRII ベクターへクローニングした。以下に TA Cloning の概略を示す。

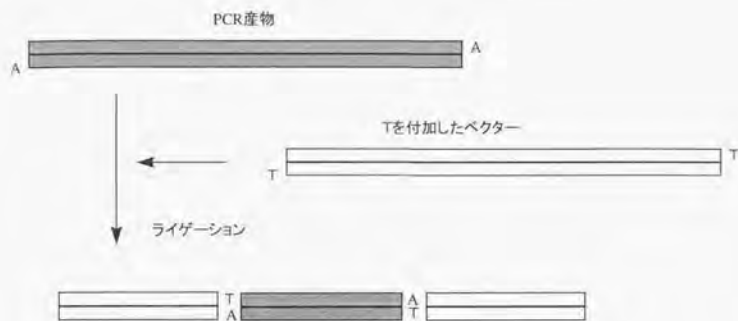


図3 TA cloning の概略

クローニングした増幅産物を含むプラスミドは、アルカリ SDS 法により精製した。これを鋳型として、M13 universe および reverse プライマーを用いて DNA シーケンスをし、その後、必要に応じてプライマーを合成しシーケンスを行った。

DNA シーケンスは、ABI373 を使い、ダイターミネーター法により行った。ダイターミネーター法によるシーケンス反応は、DNA Thermal Cycler Model 480 (Perkin-Elmer)を用いて以下の条件で行った。

表12 シーケンス反応溶液

鋳型 DNA	8 μ l (0.2~0.5 μ g)
ターミネーター反応 mix	8 μ l
プライマー	4 μ l

反応サイクル

96°C 5分 1 サイクル

↓

96°C 30 秒

50°C 15 秒

72°C 4 分

25 サイクル

3-3 結果

3-3-1 ブレークポイントの同定

ヒト・イムグロブリン H 鎖 J_H 領域をプローブとして t(11;14)転座を持つ患者の腫瘍組織細胞から抽出した DNA と正常細胞から抽出した DNA 10 μg を EcoR I、BamH I、Hind III、Pst I、Bgl II で消化した後、常法によりサザンブロッティングし、サザンハイブリダイゼーションを行った。その結果を下図に示す。

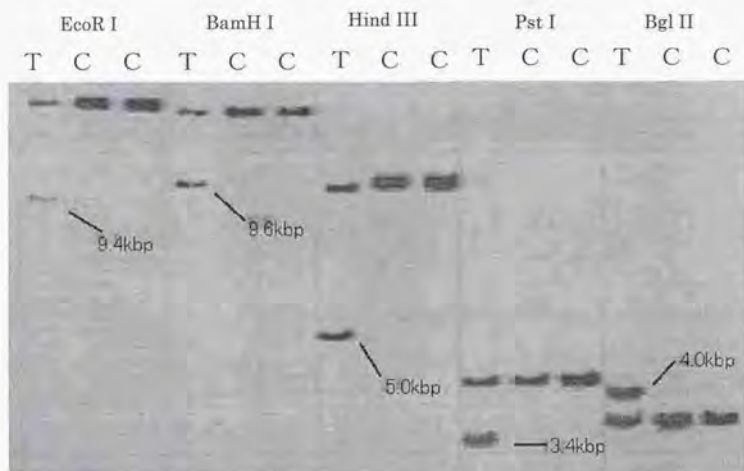


図4 breakpoint の同定 (サザンハイブリダイゼーション)

ここで T は、腫瘍細胞から抽出した DNA、C は正常細胞から抽出した DNA である。

ヒト・イムグロブリン H 鎖 J_H を用いると、腫瘍細胞に特異的なバンド (矢

印) が得られた。どの腫瘍細胞においても、正常細胞と同じ大きさのバンドと腫瘍細胞特異的なバンドの 2 つが観察され、かつシグナルの強さは、正常細胞で得られたバンドの強さの半分であった。このことは、正常細胞では、どちらの allele も同じ大きさのバンドを与え、腫瘍細胞では一方が正常と同じバンド、もう一方が腫瘍細胞特異的なバンドを与えることを示唆していると考えられる。

したがって、t(11;14)転座の breakpoint をクローニングするために、ヒト・イムノグロブリン H 鎖の J_H 領域をプローブとすることにした。

3-3-2 ブレークポイントのクローニング

スクリーニングの結果、下図に示したクローン COS10 が得られた。このクローンを用いて制限酵素地図を作製した。

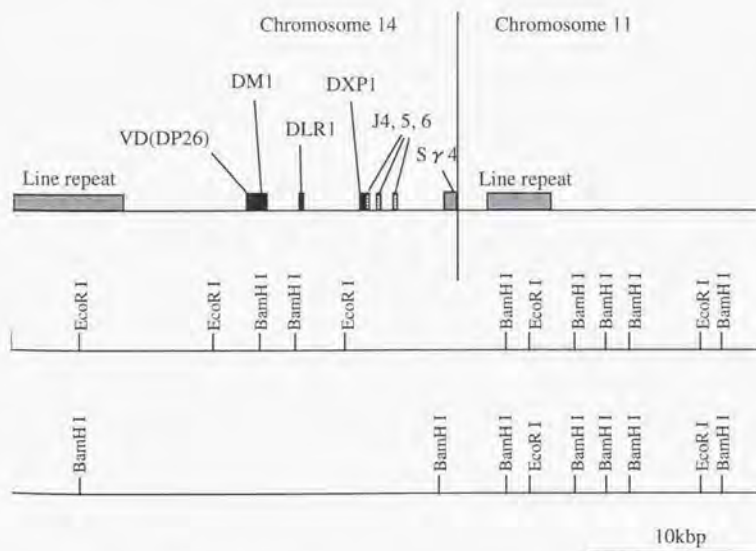


図5 breakpoint の制限酵素地図

また、「3-3-3 ブレークポイントの解析」で得られた結果と合わせて、COS10内の染色体 11 番領域をプローブとして、再度スクリーニングを行って、転座が起きていない正常な染色体 11 番 (breakpoint に隣接した領域) を含む COS6 クローンを得た (上図参照)。以下、この COS6 を用いて解析した。

3-3-3 ブレークポイントの解析

COS10 と COS6 の制限酵素地図から、breakpoint を含むと考えられる DNA 断片 (EcoR I 断片) をプラスミド pBluescript SK(-)へサブクローニングし、DNA

```
AGTTCAAATTTTCAACAAAMAGGAGGAGGGAAAGTGCTCGGTAAAAACACAACGGAGTAACTGGTTTTGGCTCCGACACAGCGGAGAGAAAGAGTGAGAAGT  CH11
|||||
AGGGGAGGGGTGCTCAGCGGGACCTGGGCTGAGCAGAGGGGATCTCAGAGGTCAGGGAGTGGTTTTGGCTCCGACACAGCGGAGGGGAGAGAAGAGTGAGAAGT  t(11;14)
|||||
AGGGGAGGGGTGGTGAAGCGGGCCCTGGGGCTGAGCAGAGGGGATGCGCTGGCTGAGGGG  CH14
```

図6 breakpoint の塩基配列

シーケンスを行い、breakpoint 周辺の塩基配列を決定した。その結果を以下に示す。

CH11、CH14 は、それぞれ正常な染色体 11 番、14 番の塩基配列である。t(11;14) は、クローニングされた EcoR I 断片から明らかにされた breakpoint の塩基配列である。

3-3-4 L P C の同定と cDNA のクローニング

エクソントラッピングの結果得られたクローンの DNA シーケンスを決定し

た。シーケンス解析から、ラビット β グロビン遺伝子配列が含まれていないものを除いて、クローン etp12 を得た。etp12 の塩基配列を下图に示す。

```
TCAGCAGGATGGACCTGGAGATGTCAGGGCTGAAGACCCCTGGAGCATGTGGCAGTGACAGTCTCCATCACTCACCCACGG  
CGCGGCAGCTTGGAGCTGAAGCTGTTCTGCCCCAGTGGCATGATGTCCTTCATCGGGCGCCCCCGCAGCATGGACTC
```

図7 etp12 の塩基配列

また、このクローンは、コスミドクローン COS6 とハイブリダイズすることから、COS6 内の配列を含んでいることが明らかになった（データは示さない）。

この etp1 の配列を用いてデータバンクを検索したところ、subtilisin serine protease との相同性が明らかになった（後述）。

完全長の cDNA を得るためにヒト Jurkat 細胞から作製した cDNA ライブラリーをスクリーニングして、完全な CDS を含むクローン pcM と 3'末端を含むクローン pcD および pc21 を得た。

データベース解析から、この遺伝子は、furin などの serine protease ファミリーに属する遺伝子であり、今まで報告されていない遺伝子であることが明らかになった。そこで、この遺伝子を Lymphoma Proprotein Convertase (LPC 遺伝子) と名付けた。以下、この論文では、LPC を呼ぶことにする。

他のファミリーとの比較は、マウス LPC の配列と共に考察の中で述べることにする。

ヒト LPC cDNA の塩基配列を以下に示す。ヒト LPC cDNA をシーケンスは、染色体 DNA をシーケンスした時と同様に、必要に応じて 17~24mer のオリゴヌクレオチドを合成した。

ほぼ同時期に、他のグループから LPC (proprotein convertase 7; PC7) のクローニングが報告された。彼らの報文によると、ヒト LPC cDNA は、ここで示した配列より長いことが明らかになっている。また、後に述べるマウス LPC cDNA

および染色体 DNA のシーケンスから明らかになったエクソン 1 とエクソン 2
の一部が、この cDNA 配列には含まれていない。

1 AGA ATC CAG TCC ACT GCT CTG ATG CCG AAG GGG AGG CAG AAA GTG
Met Pro Lys Gly Arg Gln Lys Val

46 CCA CAC TTG GAT GCC CCC CTG GGC CTG CCC ACC TGC CTC TGG CTG
Pro His Leu Asp Ala Pro Leu Gly Leu Pro Thr Cys Leu Trp Leu

91 GAA TTA GCC GGG CTC TTC TTA CTG GTT CCC TGG GTC ATG GGC CTG
Glu Leu Ala Gly Leu Phe Leu Leu Val Pro Trp Val Met Gly Leu

136 GCA GGG ACA GGT GGG CCT GAT GGC CAG GGC ACA GGG GGG GCG AGC
Ala Gly Thr Gly Gly Pro Asp Gly Gln Gly Thr Gly Gly Ala Ser

181 TGG GCT GTG CAC CTG GAA AGC CTG GAA GGT GAC GGG GAG GAA GAG
Trp Ala Val His Leu Glu Ser Leu Glu Gly Asp Gly Glu Glu Glu

226 ACT CTT GAG CAG CAG GCG GAT GCC TTG GCC CAG GCA GCA GGG CTG
Thr Leu Glu Gln Gln Ala Asp Ala Leu Ala Gln Ala Ala Gly Leu

271 GTG AAT GCT GGA CGC ATC GGA GAG CTT CAG GGG CAC TAC CTC TTT
Val Asn Ala Gly Arg Ile Gly Glu Leu Gln Gly His Tyr Leu Phe

316 GTC CAG CCT GCT GGG CAC AGG CCG GCC CTG GAG GTG GAG CCC ATC
Val Gln Pro Ala Gly His Arg Pro Ala Leu Glu Val Glu Pro Ile

361 CGC CAG CAG GTG GAG GCT GTG TTG GCT GGG CAT GAA GCT GTG CGC
Arg Gln Gln Val Glu Ala Val Leu Ala Gly His Glu Ala Val Arg

Cleavage site

406 TGG CAC TCA GAG CAG AGG CTG CTA AGG CGG GCC AAG CGC AGC GTC
Trp His Ser Glu Gln Arg Leu Leu Arg Arg Ala Lys Arg Ser Val

451 CAC TTC AAC GAC CCC AAG TAC CCG CAG CAA TGG CAC CTG AAT AAC
His Phe Asn Asp Pro Lys Tyr Pro Gln Gln Trp His Leu Asn Asn

496 CGA CGG AGC CCG GGC AGG GAC ATC AAC GTG ACG GGT GTG TGG GAA
Arg Arg Ser Pro Gly Arg Asp Ile Asn Val Thr Gly Val Trp Glu

541 CGC AAT GTG ACT GGG CGA GGG GTG ACG GTG GTG GTA GTG GAT GAC
Arg Asn Val Thr Gly Arg Gly Val Thr Val Val Val Val Asp Asp

586 GGA GTG GAA CAC ACC ATC CAG GAC ATT GCA CCC AAC TAT AGC CCT
Gly Val Glu His Thr Ile Gln Asp Ile Ala Pro Asn Tyr Ser Pro

631 GAG GGT AGC TAT GAC CTC AAC TCT AAT GAC CCT GAC CCC ATG CCC
Glu Gly Ser Tyr Asp Leu Asn Ser Asn Asp Pro Asp Pro Met Pro

676 CAC CCG GAT GTG GAG AAT GGC AAC CAC CAT GGC ACG CGA TGT GCA
His Pro Asp Val Glu Asn Gly Asn His His Gly Thr Arg Cys Ala

721 GGA GAG ATC GCG GCT GTG CCC AAC AAC AGC TTC TGT GCC GTG GGC
Gly Glu Ile Ala Ala Val Pro Asn Asn Ser Phe Cys Ala Val Gly

766 GTG GCC TAC GGG AGC CGC ATC GCA GGT ATC CGG GTA CTG GAT GGA
Val Ala Tyr Gly Ser Arg Ile Ala Gly Ile Arg Val Leu Asp Gly

- 811 CCT CTC ACA GAC AGC ATG GAG GCA GTG GCG TTC AAC AAG CAC TAT
Pro Leu Thr Asp Ser Met Glu Ala Val Ala Phe Asn Lys His Tyr
- 856 CAG ATC AAT GAC ATC TAC AGC TGC AGC TGG GGA CCA GAT GAC GAT
Gln Ile Asn Asp Ile Tyr Ser Cys Ser Trp Gly Pro Asp Asp Asp
- 901 GGG AAG ACA GTG GAT GGC CCC CAT CAG CTT GGA AAG GCT GCC TTA
Gly Lys Thr Val Asp Gly Pro His Gln Leu Gly Lys Ala Ala Leu
- 946 CAA CAT GGG GTG ATT GCT GGT CGC CAG GGC TTT GGG AGC ATC TTT
Gln His Gly Val Ile Ala Gly Arg Gln Gly Phe Gly Ser Ile Phe
- 991 GTG GTA GCC AGT GGC AAC GGA GGC CAA CAC AAC GAC AAC TGC AAC
Val Val Ala Ser Gly Asn Gly Gly Gln His Asn Asp Asn Cys Asn
- 1036 TAC GAT GGC TAC GCC AAC TCC ATC TAC ACC GTC ACC ATA GGA GCT
Tyr Asp Gly Tyr Ala Asn Ser Ile Tyr Thr Val Thr Ile Gly Ala
- 1081 GTG GAT GAG GAG GGA CGC ATG CCT TTC TAT GCA GAA GAA TGT GCC
Val Asp Glu Glu Gly Arg Met Pro Phe Tyr Ala Glu Glu Cys Ala
- 1126 TCC ATG CTG GCA GTC ACC TTC AGT GGT GGG GAC AAG ATG CTT CGG
Ser Met Leu Ala Val Thr Phe Ser Gly Gly Asp Lys Met Leu Arg
- 1171 AGC ATT GTG ACC ACT GAC TGG GAC CTT CAG AAG GGC ACT GGC TGC
Ser Ile Val Thr Thr Asp Trp Asp Leu Gln Lys Gly Thr Gly Cys

- 1216 ACT GAG GGC CAC ACA GGG ACC TCA GCT GCA GCG CCT CTG GCA GCT
Thr Glu Gly His Thr Gly Thr Ser Ala Ala Ala Pro Leu Ala Ala
- 1261 GGC ATG ATA GCC TTA ATG CTG CAG GTG CCG CCC TGC CTC ACG TGG
Gly Met Ile Ala Leu Met Leu Gln Val Arg Pro Cys Leu Thr Trp
- 1306 CGT GAC GTC CAG CAC ATC ATT GTC TTC ACA GCC ACC CGG TAT GAG
Arg Asp Val Gln His Ile Ile Val Phe Thr Ala Thr Arg Tyr Glu
- 1351 GAT CGC CGT GCA GAG TGG GTC ACC AAC GAG GCA GGC TTC AGC CAT
Asp Arg Arg Ala Glu Trp Val Thr Asn Glu Ala Gly Phe Ser His
- 1396 AGC CAC CAG CAC GGT TTC GGC CTC CTC AAC GCC TGG AGG CTC GTG
Ser His Gln His Gly Phe Gly Leu Leu Asn Ala Trp Arg Leu Val
- 1441 AAT GCA GCC AAG ATC TGG ACA TCT GTC CCT TAC TTA GCA TCC TAC
Asn Ala Ala Lys Ile Trp Thr Ser Val Pro Tyr Leu Ala Ser Tyr
- 1486 GTC AGT CCC GTG TTA AAA GAA AAC AAG GCG ATT CCG CAG TCC CCC
Val Ser Pro Val Leu Lys Glu Asn Lys Ala Ile Pro Gln Ser Pro
- 1531 CGT TCC CTG GAG GTC CTG TGG AAT GTC AGC AGG ATG GAC CTG GAG
Arg Ser Leu Glu Val Leu Trp Asn Val Ser Arg Met Asp Leu Glu
- 1576 ATG TCA GGG CTG AAG ACC CTG GAG CAT GTG GCA GTG ACA GTC TCC
Met Ser Gly Leu Lys Thr Leu Glu His Val Ala Val Thr Val Ser

1621 ATC ACT CAC CCA CGG CGC GGC AGC TTG GAG CTG AAG CTG TTC TGC
Ile Thr His Pro Arg Arg Gly Ser Leu Glu Leu Lys Leu Phe Cys

1666 CCC AGT GGC ATG ATG TCC CTC ATC GGC GCC CCC CGC AGC ATG GAC
Pro Ser Gly Met Met Ser Leu Ile Gly Ala Pro Arg Ser Met Asp

1711 TCG GAT CCC AAC GGC TTC AAT GAC TGG ACC TTC TCC ACT GTG CGA
Ser Asp Pro Asn Gly Phe Asn Asp Trp Thr Phe Ser Thr Val Arg

1756 TGC TGG GGG GAG AGA GCC CGA GGG ACC TAC AGG CTT GTC ATC AGG
Cys Trp Gly Glu Arg Ala Arg Gly Thr Tyr Arg Leu Val Ile Arg

1801 GAT GTC GGG GAT GAG TCA TTC CAG GTC GGC ATC CTC CGG CAA TGG
Asp Val Gly Asp Glu Ser Phe Gln Val Gly Ile Leu Arg Gln Trp

1846 CAG CTG ACC CTA TAT GGC TCT GTG TGG AGT GCA GTA GAC ATC AGG
Gln Leu Thr Leu Tyr Gly Ser Val Trp Ser Ala Val Asp Ile Arg

1891 GAC AGA CAA AGG CTG TTA GAG AGT GCC ATG AGT GGA AAA TAC CTG
Asp Arg Gln Arg Leu Leu Glu Ser Ala Met Ser Gly Lys Tyr Leu

1936 CAC GAT GAC TTC GCC CTG CCC TGC CCA CCG GGG CTG AAA ATT CCT
His Asp Asp Phe Ala Leu Pro Cys Pro Pro Gly Leu Lys Ile Pro

1981 GAG GAA GAT GGT TAC ACC ATC ACC CCC AAC ACC CTC AAG ACC CTG
Glu Glu Asp Gly Tyr Thr Ile Thr Pro Asn Thr Leu Lys Thr Leu

2026 GTG CTG GTA GGC TGT TTC ACC GTC TTC TGG ACT GTT TAC TAC ATG
Val Leu Val Gly Cys Phe Thr Val Phe Trp Thr Val Tyr Tyr Met

2071 CTG GAA GTA TAT TTG AGC CAG AGG AAT GTG GCT TCC AAT CAA GTT
Leu Glu Val Tyr Leu Ser Gln Arg Asn Val Ala Ser Asn Gln Val

2116 TGT AGG AGT GGA CCC TGC CAC TGG CCC CAT CGG AGC CGG AAA GCC
Cys Arg Ser Gly Pro Cys His Trp Pro His Arg Ser Arg Lys Ala

2161 AAG GAG GAA GGG ACA GAG CTA GAA TCA GTG CCA CTT TGC AGC AGC
Lys Glu Glu Gly Thr Glu Leu Glu Ser Val Pro Leu Cys Ser Ser

2206 AAG GAT CCA GAC GAA GTG GAA ACA GAG AGC AGG GGC CCT CCC ACC
Lys Asp Pro Asp Glu Val Glu Thr Glu Ser Arg Gly Pro Pro Thr

2251 ACC TCT GAC CTC CTT GCC CCA GAC CTG CTG GAG CAA GGG GAC TGG
Thr Ser Asp Leu Leu Ala Pro Asp Leu Leu Glu Gln Gly Asp Trp

2296 AGC CTG TCC CAG AAC AAG AGC GCC CTG GAC TGC CCT CAT CAG CAC
Ser Leu Ser Gln Asn Lys Ser Ala Leu Asp Cys Pro His Gln His

2341 CTA GAC GTA CCG CAC GGG AAG GAG GAG CAG ATC TGC TGA CCT CAG
Leu Asp Val Pro His Gly Lys Glu Glu Gln Ile Cys

2386 GGC CTG ACA GTG TGG GAC AGG CTC TTC TTT CCC AAA ATT AGG GAG

2431 CTC TTG ACA GAA AGC AGT TCT GAT GCT TAC ATC TGG AAT CTG AGG

2476 CAT CCT CTG ACT CCA CTC AAA GAG GGT GAG GGC CTT CTT AAG ATA
 2521 CAA ATG GTG GAG GAT TGC TGC CAG AGA AGT CTG GTC AGA GCC ACA
 2566 GGG TCT GCC TCC AGC CAA ACG GGA GCT TTT GGT GAG AAG GTG TTG
 2611 GAC AGG GGA TTG GCG CCC CCC TTT GGT TTG GCC TCC ATC CTC ATC
 2656 TCT CTT GGG CCA AGC CAG CTG CCT AGG TCC CCC AAG CAT GGG GGA
 2701 CCC CTT CCC ACA TAT AAG TTG AGA AGG TGC CTG CCA TAG CCA GGA
 2746 GCG CAT CTC AAT GGA AAC ATC ACT GGG GTC ACT TGG GAA GAG GAC
 2791 TTC GGG GTA GAG GCT GGG AGG AGC CCC TGG ACA TGC CTG TCC TGA
 2836 AAG CGG CTG CCT CCA TTA TCC ATT CCC AAG ATG CCT GAT CAG AAA
 2881 CCA ACC ATG AAT GAA CCC CTG GCT CCT TCA CCA CCC CCA CGA TTG
 2926 GTA TGA TGC TGC CGG CAC AGC TGG GAT ACA CAC GGC TCC CCC AGG
 2971 CCT GAG CTG CTT CAC TAG GGA ATC CTG CGG CAG GAC TGC AGA GCA
 3016 GAT GGC AGA TGC ACA TGT TGG AGG AGA GAG CCT TGG GAG CCA CTG
 3061 CCA CTC CAG TCC TGC CAC CAC CCT GTC TTC CTC TGC AAG TGC TGA
 3106 GGG AAA TGG CCT TCC CGC CGG AGG CCA GCT ATC TGC CTG ACA GGC
 3151 TGT GAC TCT TCT CTC AAC CTT GGC CTT CTC CCC TCT TCT GAG CTA
 3196 GTT GGT TGA ATT TTT TTT AAT GCT TAA GAT TTG TTT TTC TCT TTT
 3241 CAC AGC AAC ATT TTC TTG AAT TTT TTT CTG CAC AGC TTT TCC AAA
 3286 ATA AAA ACC TTC CAA ACA AAA AAA AAA AAA AAA AAA AAA

ポリ A 鎖付加シグナルは、下線で示した。また、proprotein convertase に保存されている活性に関与しているアミノ酸（4アミノ酸残基）は、四角で囲って示した。

Catalytic domain は下線、middle domain には破線で示した。

また、LPC が成熟化する際に切断される部位は縦線で示した。

さらに 11q23 転座で見つかった breakpoint を太い縦線で示した。

次に、LPC 遺伝子内の breakpoint を示す。

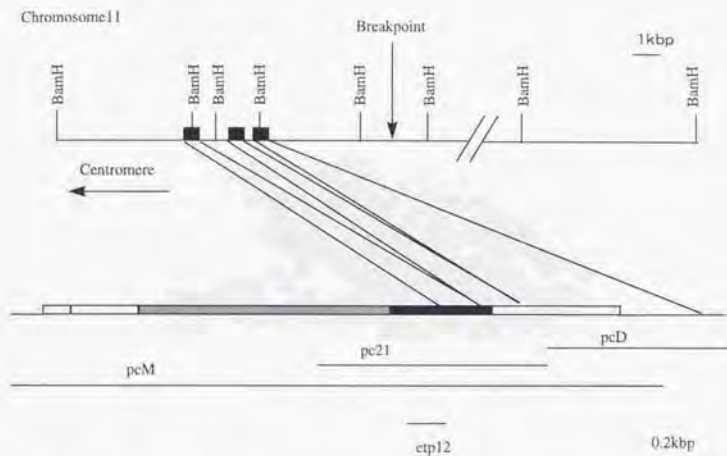


図 8 LPC 遺伝子と breakpoint

この図から明らかなように、LPC 遺伝子内の breakpoint は、coding 領域内ではなく 3' non-coding 領域であった。他の 11q23 が関与している転座では inframe な融合タンパク質ができて例のみが報告されているが、LPC 遺伝子による lymphoma 発症は、別のメカニズムであることが考えられた。

次に、ノーザンハイブリダイゼーションの結果を示す。



図 9 ノーザンハイブリダイゼーション (ヒト mRNA)

左から spleen, thymus, prostate, testis, ovary, small intestine, colon, peripheral blood
leukocyte,



図10 ノーザンハイブリダイゼーション (マウス mRNA)

左から、heart, brain, spleen, lung, liver, skeltal muscle, kedney, testis

ヒト、マウスともに調べたすべての組織において発現していることが明らかになった。

3-3-5 エクソン・イントロン構造の解析

ヒト LPC cDNA がクローニングされていたため、塩基配列情報を元にして PCR 用プライマーを作製した。これを用い、すでにクローニングされていたヒト LPC 染色体 DNA を含む 3 つの P1 クローンを鋳型として、PCR により必要な部分 (エクソン・イントロンを含む領域) のみを増幅し、DNA シーケンスを行った。

以下に、PCR によって増幅したエクソンの例を示す。

1 2 3 4 5 6 7

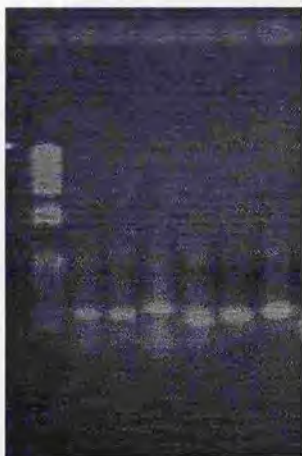


図11 エクソン領域の増幅 (PCR)

レーン1 : ϕ X174/Hae III

レーン2 : エクソン 11

レーン6 : エクソン 15

レーン 3 : エクソン 12

レーン 7 : エクソン 16

レーン 4 : エクソン 13

レーン 5 : エクソン 14

エクソン/イントロンのコンセンサス配列を以下に示す。



図12 エクソン・イントロンのコンセンサス配列

下線は、100%保存されている配列を示す。

以下に、ヒト LPC 遺伝子のエクソン・イントロン構造を示す。

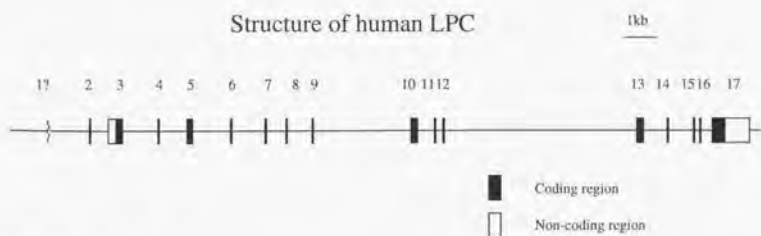


図 13 ヒト LPC 遺伝子のエクソン・イントロン構造

得られた P1 クローンを上記の PCR を用いた方法で解析したところ、クロー

ン N1 と N3 は、エクソン 2 から 17 までのすべてを含んでいたが、クローン N2 はエクソン 13 以降のみを含んでいることが明らかになった。

当初、これは、N2 がエクソン 13 以降を含んだ 3'領域のみを持ったクローンであると考えたが、詳しく調べるうちにそうではないことがわかった。

エクソン 17 の 5'側のプライマーと 3'側のプライマーを用いて PCR を行ったときに、A1、B3 では 1.2kbp のバンドが得られたのに対し、D4 ではまったくバンドが得られなかった。しかし、エクソン 17 のほぼ真ん中付近とアニーリングするように設計されたプライマーと 5'側のプライマーを用いると、どのクローンも同じ大きさのバンドを得ることができた（データは示さない）。このことは、N2 は、エクソン 17 の 3'領域が欠失していることを示唆する。また、D4 のエクソン 13 より 5'側を調べると、エクソン 12 以前のエクソンは存在しないが、少なくとも 20kbp 以上上流まではクローニングされていることが明らかになった。

以上の知見と、breakpoint の解析から得た LPC 遺伝子下流に存在する Line repeat とを合わせて考えると、以下の図のような少なくとも 2 つの LPC 遺伝子（一方はエクソン 13 からエクソン 17 の途中までのみを含む）が存在することが明らかになった。

Structure of human LPC

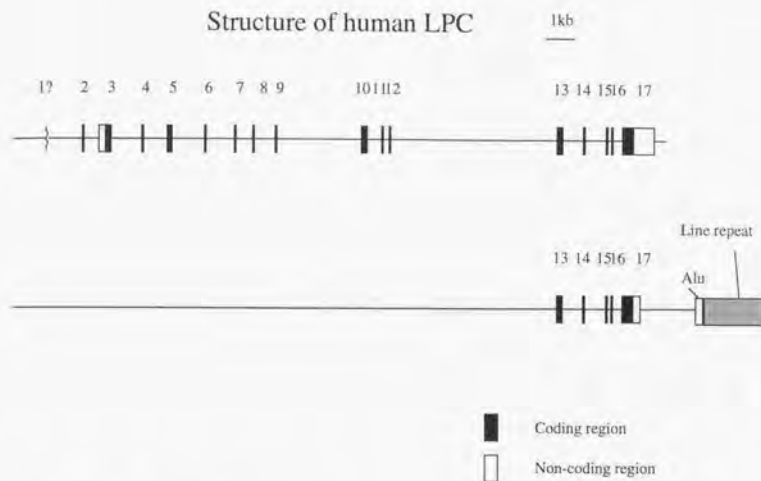


図14 ヒト LPC 遺伝子の構造

また、データは示さないが、FISH 法による詳細なマッピングによると、これらの2つの LPC 遺伝子は、11q23 内の極めて近い領域に存在することが明らかになっている。

4 マウス LPC 遺伝子のクローニングと解析

4-1 はじめに

前章で、ヒト LPC 遺伝子について解析した。この遺伝子は、リンパ腫に関与するだけでなく、タンパク質の成熟化に関与する proprotein convertase であることから、正常細胞において極めて重要な役割をになっていると考えられる。LPC 遺伝子の機能を解析するうえで、マウス LPC 遺伝子をクローニングすることは有用である。この章では、マウス LPC 染色体 DNA の構造と cDNA をクローニングしたので報告する。

4-2 材料および方法

4-2-1 マウス染色体 DNA のスクリーニング

マウス LPC 染色体 DNA を得るために使用したプローブは、次のように調製した。まず、すでにクローニングされていたヒト LPC cDNA の触媒ドメインをコードしている領域の 5'末端に相補的な合成ヌクレオチドを作製した。

5'-CTTCAACGACCCCAAGTACCC-3'

図15 ヒト LPC プライマー (M27)

このプライマーと、ベクターのクローニング部位上流の M13 プライマーを使用して、PCR を行った。これにより、ヒト LPC 遺伝子の prepro 領域のみが増幅された。

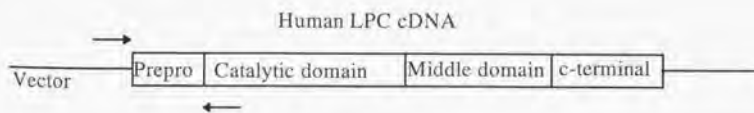


図 16 ヒト LPC cDNA の prepro 領域の増幅

PCR は、以下の条件で行った。

表13 反応溶液組成

鋳型 DNA (ヒト LPC cDNA)	10.0 μ l
10 \times PCR バッファー	5.0 μ l
25mM MgCl ₂	5.0 μ l
2mM dNTPs (each)	5.0 μ l
プライマー 1 (20mM)	2.5 μ l
プライマー 2 (20mM)	2.5 μ l
H ₂ O	19.0 μ l
1u/ μ l Taq DNA ポリメラーゼ	1.0 μ l

反応サイクル

94°C 5分

60°C 2分 1サイクル

72°C 10分

↓

94°C 1分

60°C 1分 30サイクル

72°C 1分

↓

72°C 10分 1サイクル

反応終了後、0.8%アガロースゲル電気泳動を行い、目的の増幅物 (prepro 領域) を含むアガロース片 (約 600bp) を切り出した。このアガロース片に、H₂O を加えてアガロースの濃度が 0.4%以下になるように調製した。

このようにして増幅・精製された LPC prepro 領域をプローブとして、マウス染色体 DNA に対してサザンハイブリダイゼーションを行い、プローブが他のセリンプロテアーゼ遺伝子などを認識しないことを確認した。

サザンハイブリダイゼーションは、以下のように行った。

マウス高分子量 DNA 10 μ g を 50units の BamH I, EcoR I, Hind III で消化し、

0.7%のアガロースゲル (25V、一晚) で、電気泳動した。その後、エチジウムブロマイド $0.5 \mu\text{g}/\mu\text{l}$ を含む $1\times\text{TBE}$ バッファーに 15 分間浸して染色し、泳動を確認するとともに、スケールをあてて写真を撮った。0.25N HCl 溶液に 30 分間、0.4N NaOH 溶液に 20 分間浸した後、中和することなく 0.4N NaOH 溶液を用いて、アルカリトランスファーを行った。トランスファーを 2 時間行った後、 $2\times\text{SSC}$ 溶液でメンブレンをリンスし風乾した。メンブレンは、Hybond N+ (Amersham) を用いた。

プローブの標識は、上記の増幅産物を含むアガロース片を 100°C で 10 分間加熱してとかした後、ランダムプライマーラベリングキット (Amersham) を用いて行った。標識には、約 50ng の DNA を用いた。標識後、スパンカラムにより標識された DNA プローブを精製した。

メンブレンを以下の表に示したハイブリダイゼーション溶液中で 65°C 、1 時間プレハイブリダイゼーションした後、 100°C 、5 分間加熱後氷冷した DNA プローブを加えて、 65°C で一晚ハイブリダイゼーションを行った。

表14 ハイブリダイゼーション溶液の組成

10×Denhart's
4×SSC
1% SDS
100 $\mu\text{g}/\text{ml}$ salmon sperm DNA

ハイブリダイゼーション後、次の条件でメンブレンを洗浄した。

$2\times\text{SSC}/0.1\%$ SDS 室温、5分

$2\times\text{SSC}/0.1\%$ SDS 室温、5分

$1\times\text{SSC}/0.1\%$ SDS 65°C 、10分

その後、フィルムカセット内で -70°C 、一晩露光させた。

マウス LPC 染色体 DNA のスクリーニングは、以下のように行った。

NM539 大腸菌を LB 培地中で一晩培養し、その培溶液 0.5ml を 50ml の LB 培地 (0.2% マルトースと 10mM MgSO_4 を含む) に殖菌し 3 時間、 37°C で振とう培養した。この培溶液 3ml とファージ溶液 (2.5×10^5 pfu; 129 マウス染色体ライブラリー; λ 2001) を混ぜ、室温で 20 分間、 37°C で 10 分間インキュベートした後、50ml のトップアガロースを加えてよく混ぜた後、 $245 \times 245\text{mm}$ のプレート上に広げた。これを 37°C で一晩インキュベートしてプラークを得た。

1 時間 4°C で冷やした後、次のようにハイブリダイゼーション用メンブレンを調製した。メンブレンは、Hybond N+ (Amersham) を使用した。

まず、プレート上にメンブレンを載せ、1 分間放置した。この時、メンブレン上の 3 個所に針で印をつけた。メンブレンを 3MM ペーパーの上に置き、最低 10 分間乾燥させた。その後、

0.2N NaOH/1.5M NaCl 1 分間

1.5M Tris-HCl/2 \times SSC 1 分間

2 \times SSC 1 分間

処理し、3MM ペーパー上で乾燥させた。

LPC prepro 領域をプローブとして、129 マウスの染色体 DNA ライブラリーをスクリーニングした。プローブの標識は、サザンハイブリダイゼーションと同様に行った。

メンブレンを以下の表に示したハイブリダイゼーション溶液中で 42°C 、1 時間プレハイブリダイゼーションした後、 100°C 、5 分間加熱後氷冷した DNA プローブを加えて、 42°C で一晩ハイブリダイゼーションを行った。

表15 ハイブリダイゼーション溶液の組成

50% Formamide
5×Denhart's
5×SSC
0.1% SDS
100 μ g/ml salmon sperm DNA

ハイブリダイゼーション後、次の条件でメンブレンを洗浄した。

2×SSC/0.1% SDS 室温、5分

2×SSC/0.1% SDS 室温、5分

0.2×SSC/0.1% SDS 42°C、15分

0.2×SSC/0.1% SDS 42°C、15分

その後、フィルムカセット内で-70°C、一晚露光させた。

4-2-2 欠失変異株の作成

欠失変異株は、Deletion kit (宝酒造) を用いて次のように作成した。

まず、欠失変異株取得に使用する Plasmid DNA を CsCl 密度勾配法により調製した。この plasmid DNA 10g を Kpn I および Sal I で完全に消化した。フェノール抽出・エタノール沈殿により精製した後、100 μ l の exonuclease III buffer で溶解した。これを 37°C に温めた後、exonuclease III を 1 μ l 加え、軽く混ぜてすばやく 37°C に戻した。45 秒から 1 分毎に 10 μ l ずつを別のチューブ (あらかじめ Mung Bean nuclease buffer を 100 μ l 分注しておく) に移した。すべての溶液を移し終えた後、65°C で 5 分間インキュベートして、exonuclease III を失活させた。チューブを 37°C に戻した後、2 μ l の Mung Bean nuclease を加え軽く混ぜて、37°C で 45~60 分間インキュベートした。フェノール抽出・エタノール沈

殿で精製した後、50 μ l の Klenow buffer に溶かし、1 μ l の Klenow fragment を加えて、37 $^{\circ}$ C で 15 分間インキュベートした。この溶液 10 μ l を別のチューブに移し、ライゲーション溶液 A を 100 μ l、溶液 B を 12 μ l 加えて、16 $^{\circ}$ C で 1 時間インキュベートした。その後、フェノール抽出・エタノール沈殿で精製し、40 μ l の反応溶液 (10units の制限酵素を含む) で 1 時間消化した。この溶液をそのまま用いて、定法により大腸菌を形質転換した。大腸菌は XL-1 Blue、あるいは DH5 α を用いた。得られたコロニー 50 個を 5ml の LB 液体培地で 37 $^{\circ}$ C、一晩振とう培養して、アルカリ SDS 法により plasmid DNA を精製した。

4-2-3 DNA シーケンス

2-2-1 で得られたファージ DNA を Xho I または Xba I (入マルチクローニングサイト) によって断片化し、pBluescript SK(+)へサブクローニングした。サブクローニング後、それでも大きな断片については HindIII で消化し再度サブクローニングを行い、1~4kbp 程度の断片とした。

さらにターゲットとなる配列を含んでいると考えられる領域、および 3kbp 以上の断片に関しては、欠変異株を作成した後 DNA シーケンスを行った (4-2-2 欠変異株の作成参照)。

DNA シーケンスは、ABI373A を使い、ダイターミネーター法により行った。ダイターミネーター法によるシーケンス反応は、DNA Thermal Cycler Model 480 (Perkin-Elmer) を用いて以下の条件で行った。

表16 シーケンス反応溶液

鋳型 DNA	8 μ l (0.2~0.5 μ g)
ターミネーター反応 mix	8 μ l
プライマー	4 μ l

反応サイクル

96°C 5分 1 サイクル

↓

96°C 30秒

50°C 15秒 25 サイクル

72°C 4分

シーケンスに用いたプライマーは、M13 forward と M13 reverse、それに得られたシーケンスからさらにその下流をシーケンスするためのプライマーを調製した。

4-2-4 エクソン・イントロン構造の解析

エクソン・イントロンの解析は、次のように行った。まず、得られた DNA 配列が LPC 遺伝子を含んでいること、さらに、他のセリンプロテアーゼ遺伝子ではないことを確認するために、Blastn プログラムを使用して DNA データバンクを検索した。

検索結果を元に、マウス染色体配列中のエクソン領域を推定し、その領域を含む前後 1kb ほどの配列を、Gene Finder の HSPL (search for potential splice sites) により解析しエクソンを同定した。その後、スプライシング部位に保存されている配列、およびブランチ部位に保存されている配列を確認した。

また、マウス LPC cDNA 配列を決定 (4-2-5 参照) した後は、マウス LPC cDNA 配列を元にして PCR プライマーを作製し、鋳型 DNA としてスクリーニングしたマウス LPC を含むファージ DNA、およびクローニングできなかった領域に関してはマウス染色体 DNA を使用した。PCR は以下の条件で行った。

表17 反応溶液組成

鋳型 DNA	10.0 μ l
10 \times PCR バッファー	5.0 μ l
25mM MgCl ₂	5.0 μ l
2mM dNTPs (each)	5.0 μ l
プライマー 1 (20mM)	2.5 μ l
プライマー 2 (20mM)	2.5 μ l
H ₂ O	19.0 μ l
1u/ μ l Taq DNA ポリメラーゼ	1.0 μ l

反応サイクル

94°C 5分

↓

94°C 30秒

60°C 30秒 30 サイクル

72°C 1~10分

反応終了後、1.4%アガロースゲル電気泳動にて増幅された領域にイントロンが含まれているかどうかを確認した。

また、増幅した PCR 産物は、Geneclean II (BIO101)を用いて精製した後、TA-cloning kit (Invitrogen)を用いてプラスミドヘクローニングし、DNA シーケンスを行った。TA-cloning の概略を以下の図に示す。

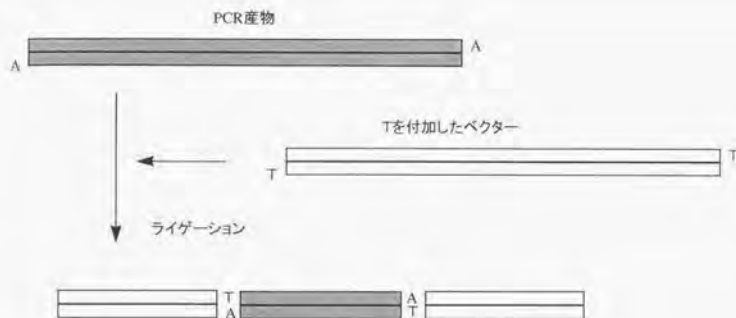


図17 TA-cloning の概略

Taq polymerase を用いて PCR 反応を行うと、増幅産物の 3'末端にアデニン(A)が1塩基だけ付加される。この A と相補的な T を付加したベクターを用いてプラスミドベクターへのクローニングを行うと、 blunt end ・ライゲーションよりクローニング効率が上がることが知られている。

PCR によるエクソン・イントロン構造決定法の概略を以下に示す。

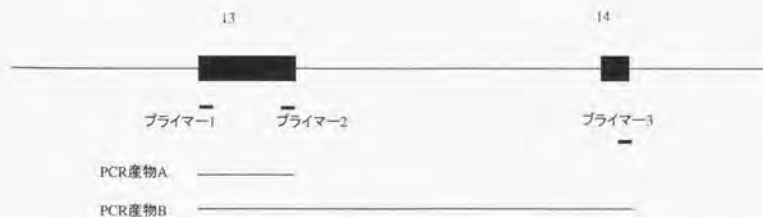


図18 PCR によるエクソン・イントロン構造の決定

プライマー 1 と 3 による増幅産物 B は、エクソン 13、イントロン 13、そしてエクソン 14 と含んでいると考えられる。一方、プライマー 1 と 2 による増

増幅産物 A は、エクソン 13 のみを含んでいる。したがって、増幅産物 B の長さから増幅産物 A の長さを引くと、イントロン 13 (エクソン 14 の一部も含む) を求めることができる。

4-2-5 マウス cDNA のスクリーニング

マウス LPC cDNA は RACE 法により取得した。

RACE 法に用いた mRNA は、マウスリンパ腫細胞 EL4 (約 10^7 細胞) より、Micro-FastTrack (Invitrogen) を用いて調製した。

また、すでにエクソン・イントロン構造の解析により明らかになったエクソンの配列から、RACE 法に用いる次の 2 つのオリゴヌクレオチドを合成した。

MA5 5'-CCGGGTGCTGGATGGACCCTTACTGACA-3'

MA3 5'-GCACCTGCAGCATGAGGGCTATCATTCCA-3'

図19 RACE 用プライマー

鋳型となる cDNA は以下のように調製した。

$1\mu\text{l}$ の $1\mu\text{g}/\mu\text{l}$ mRNA と $1\mu\text{l}$ の $10\mu\text{M}$ の cDNA 合成プライマー (5'-TTCTAGAATTCAGCGGCCGC(T)₃₀N₁N-3')、 $3\mu\text{l}$ の H₂O を加えて 70°C で 2 分間インキュベートした。その後、以下の試薬を加えて 42°C で 1 時間 30 分インキュベートした。

表18 First strand synthesis reaction

5 μ l	mRNA 溶液 (含 primer)
2 μ l	5 \times first strand buffer
1 μ l	10mM dNTP mix
1 μ l	100units/ μ l MMLV reverse transcriptase

First strand buffer : 250mM Tris (pH8.0)/30mM MgCl₂/375mM KCl

その後、直ちに以下の試薬を加えて、16°Cで 1 時間 30 分間インキュベートした。

表19 Second strand synthesis reaction

10 μ l	First strand reaction
48.4 μ l	H ₂ O
16 μ l	5 \times Second strand buffer
1.6 μ l	10mM dNTP mix
4 μ l	Second strand enzymes

5 \times Second strand buffer : 500mM KCl/50mM (NH₄)₂SO₄/MgCl₂/0.75mM β -NAD/Tris (pH7.5)/0.25mg/ml BSA

Second strand enzymes : 6units/ μ l E. coli DNA polymerase I/1.2units/ μ l E. coli DNA ligase/0.25units/ μ l E. coli Rnase H

さらに 2 μ l の 10units/ μ l T4 DNA polymerase を加えて、16°Cで 45 分間インキュベートした。その後、4 μ l の 0.2M EDTA/2mg/ml glycogen を加えて反応を停止させた。100 μ l のフェノール/クロロホルム/イソアミルアルコール(25:24:1)で 1 度、クロロホルム/イソアミルアルコール(24:1)で 1 度、抽出した後、40 μ l の 4M CH₃COONH₄ と 320 μ l の 99%エタノールを加えて室温で 20 分間遠心した。上清を捨てた後、80%エタノールでリンスし乾燥させてから 10 μ l の H₂O に溶解させた。

cDNA 末端に結合させるアダプターは、以下の配列のものを用意した。

```

    _____    _____
     T7 promoter      Not I
5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCCGGGCAGGT-3'
   cDNA adaptor                                3'-CCCGTCCA-5'

```

AP1 primer

5'-CCTAATACGACTCACTATAGGGC-3'

図20 cDNA アダプターと AP1 アダプター

合成した ds cDNA に以下の試薬を加えて、16°Cで一晩インキュベートした。

表20 cDNA アダプターの連結

5 μ l	ds cDNA
2 μ l	10 μ M cDNA アダプター
2 μ l	5 \times DNA ligation buffer
1 μ l	1unit/ μ l T4 DNA ligase

反応後、70°Cで5分間インキュベートし ligase を失活させた。アダプターを連結させた ds cDNA を用いて、以下のように 3'および 5'RACE 反応を行った。

表21 5'および3'RACE 反応

	5'RACE 反応	3'RACE 反応
10×PCR buffer	5 μ l	5 μ l
H ₂ O	36 μ l	36 μ l
10mM dNTP	1 μ l	1 μ l
Klen Taq polymerase	1 μ l	1 μ l
ds cDNA (1/50 希釈)	5 μ l	5 μ l
API primer	1 μ l	1 μ l
MA3	1 μ l	—
MA5	—	1 μ l

RACE-PCR 反応は、perkin-Elmer 社の DNA Thermal Cycler 480 を用い、以下の条件で行った。

94°C 1分

↓

94°C 30秒 30 サイクル

68°C 10分

反応終了後、アガロースゲル電気泳動で確認し、目的のバンドを切り出し、Geneclean II (BIO101) を用いて精製した。その後、TA-Cloning kit (Invitrogen) を用いてプラスミドへクローニングした。クローニングした 5'および 3'RACE 産物は、ABI373 を用いダイターミネーター法により DNA シーケンスを行った。

得られた配列をもとに、cDNA の 5'および 3'にプライマー (MA5-2 と MA3-3) を合成 (下図) した。

MA5-2 GCGGCGGCAGCAGCAGCAGCGACAGTAGCA

MA3-3 GGAGACCTTACTTCCGGTAAAGGGACTCGTGCC

図21 マウス LPC cDNA 増幅用プライマー

すでに合成した ds cDNA を鋳型にして、以下の条件により PCR を行った。

表 22 反応溶液組成

鋳型 DNA (ds cDNA)	10.0 μ l
10 \times PCR バッファー	5.0 μ l
25mM MgCl ₂	5.0 μ l
2mM dNTPs (each)	5.0 μ l
プライマー 1 (20mM)	2.5 μ l
プライマー 2 (20mM)	2.5 μ l
H ₂ O	19.0 μ l
1u/ μ l Taq DNA ポリメラーゼ	1.0 μ l

反応サイクル

94 $^{\circ}$ C 5分

60 $^{\circ}$ C 2分 1 サイクル

72 $^{\circ}$ C 10分

↓

94 $^{\circ}$ C 1分

60 $^{\circ}$ C 1分 30 サイクル

72 $^{\circ}$ C 4分

↓

72 $^{\circ}$ C 10分 1 サイクル

このようにして得られた、ほぼ完全長のマウス cDNA は、5'および 3'RACE 産物同様に、GeneClean II (BIO101)、TA-Cloning kit (Invitrogen) を利用してプラスミドへとクローニングした。

得られたマウス cDNA クローンは、373A DNA シーケンサー (ABI)、およびダイターミネーター・サイクルシーケンスキット (ABI) を用いて DNA シーケンスを行った。シーケンスに用いたプライマーは、M13 forward と M13 reverse、それに得られたシーケンスからさらにその下流をシーケンスするためのプライマーを調製した。シーケンス反応は、「4-2-3 DNA シーケンス」と同様に行った。

4-3 結果

4-3-1 マウス染色体 DNA のスクリーニング

ヒト LPC cDNA の pro 領域をプローブにして、サザンハイブリダイゼーションを行った結果を示す。染色体 DNA は、マウス DNA $10\mu\text{g}$ を EcoR I、Hind III、ヒト DNA $10\mu\text{g}$ を EcoR I で消化したものを用いた。

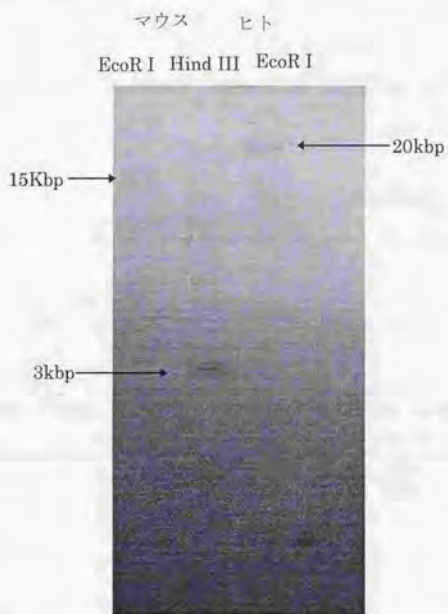


図22 pro 領域をプローブに用いたサザンハイブリダイゼーション

マウス LPC 染色体 DNA をクローニングするために用いたライブラリーは、

ICRF が提供する DNA ライブラリーサービスから供与された λ 2001 ベクターを使ったマウス染色体ライブラリーである。この染色体ライブラリーは、マウス 129 から調製した染色体 DNA を元に作られたものである。 λ 2001 のクローニング部位は BamH I であり、挿入 DNA は Sau3A I による部分消化によって調製された。

以下に λ 2001 のマップを示す。

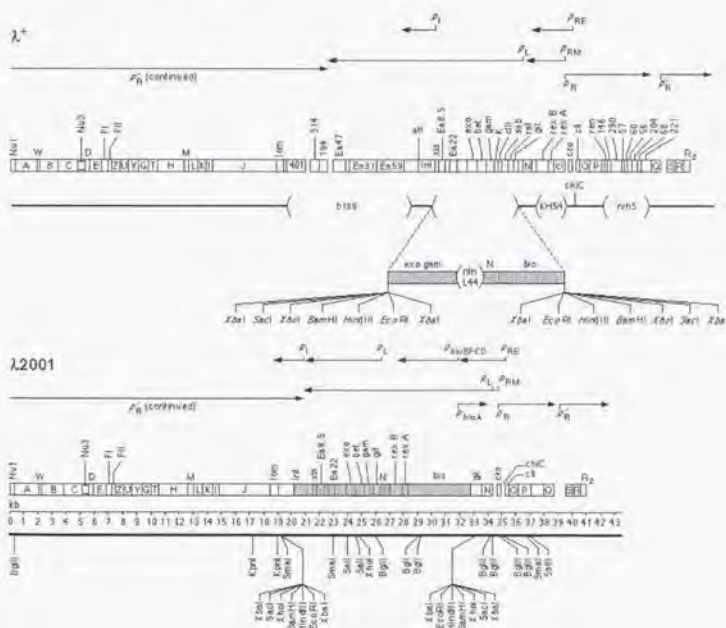


図23 λ 2001 ベクター

スクリーニングに必要なクローン数は、以下のように計算した。

目的のクローンを 99%の確率で得るためにスクリーニングしなければならない数 N は、

$$N = \frac{\ln(1-P)}{\ln(1-\frac{I}{G})}$$

与えられる。ここで、 P は目的のクローンが得られる確率、 G はマウスゲノムの大きさ、 I は λ フェージの平均のインサート長である。

$P=0.99$ 、 $G=2.7 \times 10^9$ 、 $I=15\text{Kbp}$ とすると

$$N = \frac{\ln(1-P)}{\ln(1-\frac{I}{G})} = \frac{\ln(1-0.99)}{\ln(1-\frac{15 \times 10^4}{2.7 \times 10^9})} \cong 7 \times 10^5$$

となる。したがって、約 1×10^6 クローンをスクリーニングすればよいということである。

245×245cm プレートには、約 2.5×10^8 pfu のフェージクローンをプレーティングすることができるので、最低 4 枚のプレートをスクリーニングすることにした。

しかし、4 枚のプレートをスクリーニングした結果、1 次スクリーニングで 6 クローン、2 次スクリーニングで 1 クローン、最終的に 1 クローンしか得られなかったため、再度 4 プレートをスクリーニングし、合計 3 クローンを得た。以下の表にスクリーニングの結果を示す。

表23 マウス染色体 DNA のスクリーニング (5'末端)

	1次スクリーニング	2次スクリーニング	最終スクリーニング
1回目	6	1	1
2回目	6	2	2
合計	12	3	3

得られた3クローン (C6, C8, C10) をエファージのマルチクローニング部位である Xba I と Xho I で消化して、得られた DNA 断片を pBluescript SK(-)へサブクローニングした。このようにして得られた DNA 断片を元に、制限酵素地図を作製した。このサブクローンは、DNA シーケンスにも用いられた。以下に得られた3クローンの制限酵素地図と contig マップを示す。

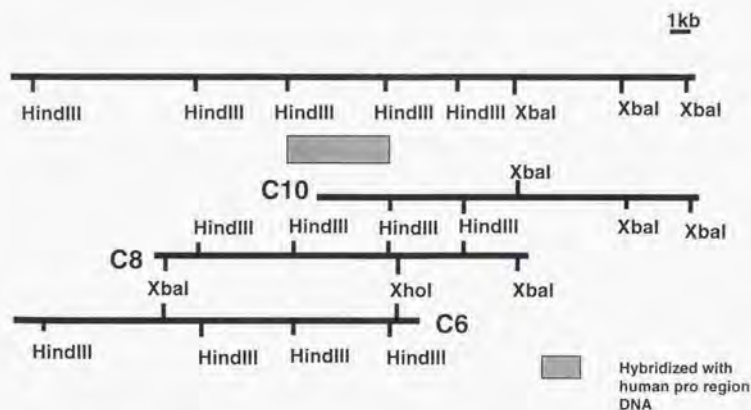


図24 マウス LPC 染色体の制限酵素地図 (5'領域)

上記2回のスクリーニングに用いたプローブがヒト LPC cDNA の5'末端から500bpの領域であったため、得られた3クローンがどれも3'領域のみを含むクローンであった。そのため、マウス LPC cDNA が得られた後に、BamH I から

下流 (1868-3390bp) の断片をプローブとして、再度 8 プレート (4 プレート×2 回) をスクリーニングした。結果を以下の表に示す。

表24 マウス染色体 DNA のスクリーニング (3'末端)

	1次スクリーニング	2次スクリーニング	最終スクリーニング
1回目	8	1	1
2回目	12	3	3
合計	20	4	4

マウス LPC 染色体の 3'領域をスクリーニングした後、サブクローニング、および制限酵素地図の作成は行わず、「3-3-3 DNA シーケンス」に示す方法により、直接必要な部分の PCR によるクローニングおよびシーケンスを行った (3-3-3 DNA シーケンス参照)。サブクローニングしなかったのは、既にマウス LPC cDNA がクローニングされていたためである。

4-3-2 欠失変異株の作成

Subcloned Fragments

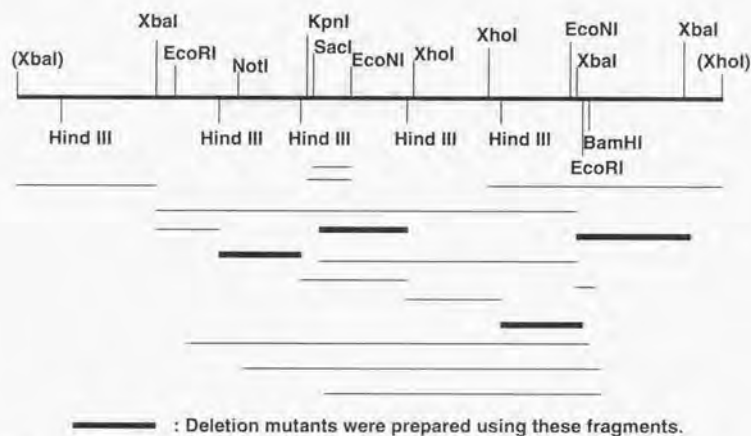


図25 サブクローニングしたマウス LPC5領域の DNA 断片

太線で示した断片は、シーケンスのために欠失変異株を取得したことを示す。

以下に、作製した欠失変異株の 1 例を示す。得られた欠失変異株をユニークな制限酵素で消化した後、アガロースゲル電気泳動したものである。



図26 シーケンスのための欠失変異株

ここで示した欠失変異株は、全体の大きさが 5.9kbp (挿入 DNA の大きさが 2.9kbp、ベクターの大きさは 3.0kbp) のクローンである。ほぼ全長にわたって、欠失変異した株が得られたことがわかる。

上記のクローンを用いて 2.9kbp の挿入 DNA をシーケンスし、contig マップを作製した。その結果を下図に示す。

欠失変異株は 2.9kbp の挿入 DNA 全体にわたって作成できたが、contig マップからわかるように、2.9kbp の中心付近ですべてのクローンのシーケンス反応が止まってしまい、それ以上シーケンスすることができなかった。この領域は、極めて G+C 含量が高く、Taq DNA polymerase を用いてはシーケンスできないと考えられる。しかし、Vent DNA polymerase を用い、かつアニーリング温度を上げて、この領域のシーケンスを得ることはできなかった。

この領域は、LPC のプロモーター配列を含む領域であり、CpG island であると考えられる。

Deletion Mutants

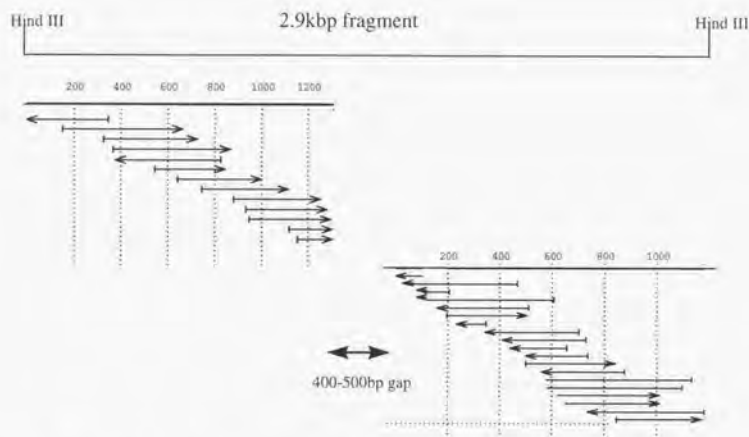


図27 欠失変異株によるシーケンスと contig

これらの contig マップを作製するためには、DNASTAR Inc.の Seqman プログラムを用いた。

4-3-3 DNA シーケンス

「4-2-1 マウス染色体 DNA のスクリーニング」および「4-2-5 マウス cDNA のスクリーニング」で得られたマウス LPC 染色体 DNA とマウス LPC cDNA クローンは、すべて ABI 社 373A DNA シーケンサー、およびダイターミネーター法により DNA シーケンスを行った。

<マウス LPC 染色体 DNA シーケンス>

マウス LPC 染色体 DNA の 5'領域(エクソン 1~エクソン 12)は、ほぼ全長に渡って DNA シーケンスを行った。3'領域については後述する。

シーケンスのためにサブクローニングした DNA 断片を以下の図に示す。

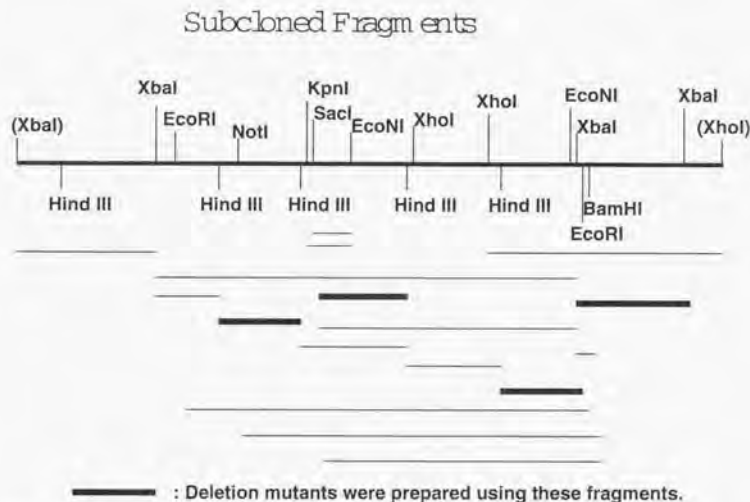


図28 サブクローニングしたマウス LPC5'領域の DNA 断片

これらのクローンをシーケンスするとともに、必要に応じて 17~24mer のオリゴヌクレオチドを合成し、プロモーター領域を含む、マウス LPC 染色体 DNA の 5'領域のシーケンスを決定した。

シーケンスの途中、2カ所シーケンスできなかった領域があるが、それは図中では nnnnn で表示した。1322、11493bp の 2カ所である。特に、最初の 1322bp からの領域は、GC 含量が 80~90% と非常に高く、欠失変異株や Vent DNA polymerase のような耐熱性がより高い酵素を使用することによって反応温度をあげてもシーケンスを決定することができなかった。この領域は、LPC 遺伝子

のプロモーター領域であり、GpC island と考えられる（後述）。

また、エクソンを同定することが目的であったため、エクソンとその近傍以外の領域には不確定の塩基が残っている。

そのシーケンスを以下に示す。

マウス LPC 染色体 DNA 配列 (エクソン1~エクソン12)

CTTARACTACTAGGCGGCAATTGGGTACCGGGCCCCCTCGAGGTCGACGGTATCG 60
 ATAAGCTTTCCTCTGCATAATGCCACTGTGTGTCAGTTCCTCCAAACACCTAAGAAAGAGT 120
 ATTCTGAAGGAGGGCAATAAAAGAGCCACTTCCGGCCATAAGTCCCAGTCAACACTGCC 180
 TCAGGCGCTGCATACACATAAAACATAAGGCAGACTACTTTTGTCTGGACAGAAAATGCA 240
 GTCATTCCAAACAGAGGATAATCTGATAAGAATTTCTGAATGGTCCACATACCTACCTC 300
 CATTTATTAAGTTGACCTTTTTGGCCAAGATTCAATAATATGTTAAATCCCTTTTGGC 360
 TGTTCAGTCAAAGCATTCTCTATTAATTAAGTGATAGGGTTTATGTGAGGTGTTTCCA 420
 AGAACAAGTCAACCCGGACCATCTCTDGGGGTGTCTTTAGCCTGACAGTATTCCTTTC 480
 CTACAGATGGGAATGTCTGGTATCTGCCAACGCCCACTAACAAAGTTACCATTGAGTCAC 540
 AGGCTACTACTCAGTACTCCAGGGAGAAAAGGGAGATGTTGAAAAACATGGTAACTCAA 600
 GTAAGGGAACCCAAGAAGACGGTTTTTTACATTTTATTTTAAAGGCATCTTTTTTGCT 660
 GCAAAATAAACTCGGCCTTCTATGGTAACAAGAAAACTAGGTGGTAAAAAGTAGGATTAG 720
 TTCCCATCTCTGCTTTTATTTTGACAGCATCCCATGAATAGAATATCTCTGAACCCAGAA 780
 GAATGCACAAGAGCAGATACGAGTCAGGCTGAATCAATAGTCACAGCTCAGTGCCTCAGT 840
 TTCCTCTCATAAGAAAAAGCTCCTAGGTGAAAAATGGGCCCTGAAACGCCATTATTGCATC 900
 TCSGATGCAATCAGTCTGCCTACAATAGGGTCGAGATCATTGCCCTCCAACAGGGAACAAG 960
 GTCTGGGTCACCGATGCCAGTAAACGAAAACCTTAGGGCAAGGGTCATACGTGAAACTC 1020
 CGCAGCTAGAGCCCCAAACGAGAGGGATGGGGGGTGGGGTCGCACATCTGGCTCAAGAC 1080
 CCAGTGTGCCCAAGGGAATCCCGCAAGGCAGTGTGGTCACCAGTCAGGCGGCGACGCAG 1140
 AAGGGACAATGGGGGAAGCTGCGCCGGTGGTCACCGCCACTTTCACAGGGAGCTGC 1200
 CGCTACGCTGGCCCTTGGCACTGGCCAGCGCCACCTCCACCCCTTCCCCTCCGTCTCAC 1260
 CGGCCGCGAGCTCCCGTCCCACTCCAGGCCGGGCTCCCGCCCTGTCCCGCCGCCCCCT 1320
 CNNNNNNNNNCGACCGCCACCGGCGCCCTCTGGCGGCAGGXGCGGACAGCGACCCGR 1380
 ACAAMGGAGCCGRTTGAGGGACCAGCCCTKAGACCCAAGCGRAGCCCGASSAATWCGCG 1440
 GGCAGAGAGACGGCGCCGGATGGAGGCAGACCTTTTCTGAGGAGTTTATGGGGAAGA 1500

GTC AAGW TATGGCAKTCGGACGGAGAAGAGCCGCGAGGGGTGTTTAGAGTTCGGAACCG 1560
 GGATGGTCTTACAGAGGACCCATGAAGGACGAGGCCCTGTAGCGGAGGGGTCTCTGT 1620
 GGCGGAGGACACGAGGATTTGAGGCGCCGGAACCTCCACAGCCCTCACGGGCCTGTGACA 1680
 GAAGCCGTCGGCGGGCGTCGGCGGAGCGCCGGAGCTCGGCCCGCTGGTCAAGAGGCCGC 1740
 AGGCCAGCCGCTGATTGGCTGCGGCGCCGGGCGGAAGTGATGTCGCTGTCAGTGGCGCC 1800
 TCCGGCTCCGGGAAGCCGAGAGTTTCCGTAGGCTGGGGCTTCGGCGGCAGCAGCAGCA 1860 Exon 1
CCGACAGTAGCAACAGAGGCTCTGGGTGAGTCGGGAAGAAATGTAGTGGAGGCTGGACCT 1920
 CGGATATACGGGGGACCATGGACCGGGCTGGTCGGGCTGCGGAGGAGGTCTCTCGCT 1980
 GAGGGACCTTGGGTCTGGCGTGGGCGACCCAGAGGGACCGGTCTCCGGTCTGATG 2040
 TTCCGGACAGCATCGGTCTTAAGGAGAGAGAGAGGCACATTGCACGGAGCAGCTGAGTCC 2100
 GCGTGGGAGTGGGGCATCAGGAAATTCACCTGTACAGACCCTGGTTCTGTGAGTGGG 2160
 GAGACGGGCTTGAAGGACGGTGGTTAAGAGCAGAACTTGAGCAGGAAAGGTGCCCGT 2220
 CTACCTTAGTTGAGGGTGATGGGAGTCTATTTTCAGAAAGTGCTTAAGATTTCTTGCTCA 2280
 GAGGTCTGAGAATAAGTGCTGCTGTGCCCTGGACCACAGGGAGAAAAAACAALCAAAC 2340
 AAACAAAAACAAAACAAAACAAAACCTCGGTTTCTCTCCAGAAATBTCTGTTGGAACCTA 2400
 CAGAGATTGTAGGAGCTGCCTGAGTGGATGGCTGCTATCAAACCAGACTTTGGATTAC 2460
 TTATATGCAGACTAGGGAAGACACCTGGGTATCTTGGGCCAAATTGAAGCAGCCCCAGGG 2520
 GCAGGAAGAATGTAATAGGAAATATGCCATGGAACCTAAGGACATGGTGGCCACACCTA 2580
 TAATAACAGCAGTAAAGGGTGGAAGCAAGAGTCTCAGCCAACCTTTATTATATAAAGA 2640
 AAGCTTGTTTAAGCCGRATGTTGGTGGCACACACCTTTGTTATCCAGCACTCTGGAAGC 2700
 GGAAGCAGGCTGGCTCTCTTGTAGTCCAGGCCAGCCTGGTATACAGAGGAAGGGGAGGTC 2760
 CAAGGAGTCTTTGGTAAAACAGTCTTTCCAGAATGTTTTAAGTCGGTACCATCCAAGTGT 2820
 ATGTCTGGACACCAAGACTGCTGCAGTGTGGGCTTGGTAGACAGTAGTGAGTTTTAAATG 2880
 AGCTAAAGACCCACATGTGCCCTTGTCTGCCCCACAGTTGCTCGAAAGTCTAATTT 2940
 GGTGGGGTTTTTTTCCCTTTGAGTGGTTGAAGACAGGGCTGGAGCCTGTGCCCTCACTG 3000
 AAAATCCAACCTTGGTGAGACGGGGCCAGCTCCTTAGTCCTTCTGCCAGTCATTTCTCT 3060
 CTCCCAGTTACCACATGCACTTATACTACATTTTATTTCCATTATCTGGGTTTCTCCCT 3120

GATTGTCCCATCTGGACTTCAAGGTGCTGGATTTCCTGGTAATGGGAGGTGCTAACAGTT	3180
AGGCTTGTCTAAGCCCTTAAAGAAGCAAAATGAAAATAACATGATAGGTTAAGAAGTGGGG	3240
TATCTGGGTTCTGTATTGGTGTGGXCTTATTGTGATGAATGTCTCCTAGTCTGTGGCCTG	3300
TTTTAGATCCCTGGAAATGACATTTAACTGGTCACTGGTCAGGCCTGGCCTCACTTT	3360
AGTTTCTGCTTTGCTGGCTTTAGTGGGACACCTTTTTTGTGTTGTTTCTGTACTTGGGGT	3420
CCACAGTCTGCAGCTGTGTTCCAGTCGACACTCACTGCTCACAGTGAAGGTGACATCT	3480 Exon 2
CCCTCCATTTTCACAAGATGAGTGTGGTATGGTTACCCAGGCCAGGACCTAAACTTCACA	3540
GAGAGCCCAAGGTAGGAATGGCTCCACAATGGGAGATAGAAGTCCACTGGGAAAAAGACTT	3600
ATGTTGTCTGTAATAAGARGGAGATAGAGCTGGGTGTTGATGGCTCATGCCTGTAATCC	3660
CAGGTAGACACAGGTGGATCTCAGTGAATTTGAGGCCAGCCTGATGGTCTACATAGTGAG	3720
TTCCAGGACAGCCAGAGAGACTACATAGTGAGATCATGTTTTTAAAAAAGAAGGAGGAA	3780
GAGAAGGAGGAGGAGGAGGAGGGTCTGATGTTGTCAGTCACTGGGAGTGATATGFACTT	3840
GGAGTAACTGAATGAGCAGCCAAAGCAAGCACCTGGATGATGCCAGGCTATTGGGGGCA	3900
CCAAGGGATGGCCATTGTGCTGTAGGAGTCACTCCACTGGGGTTAGCCCTGGCTCATACT	3960
CTACCTGGGGTTAACATTTTGTCTTAGGGACAAATGTACTIONGGTAGGGAGGGAAAGAGA	4020
CTATGAATATCAAGTACTAACCACCTTGAAAGAGGCGTGTGGGCTTCCAAGGGGAACA	4080
CAGTATGGACCTAGTCACTCTCTGAGITTCCTATAGCTGCTGGCCTTTTTTTGCTGATT	4140
CTTTCCCTTTGTTTGACCTGCTGTTCTGATGCCGAAAGGGAGGCAGAAAGTCCCACA	4200 Exon 3
CTTGGATGCCACCTGGGCTGCCATCTGCCTCTGGCTGGAATTAGCCATCTTCTTTCT	4260
GGTCCCCAGGTGATGGGCTATCAGAGGCAGGTGGGCTTGACATCTTGGGCACAGGGGG	4320
GCTGAGCTGGGCGGTACATCTGGACAGCCTAGAAGGTGAGAGGAAGGAAGAGAGTCTGAC	4380
GCAACAGGCCGATGCTGTGGCCAGGCAGCAGGGCTTGTGAATGCTGGGCGCATTGGAGA	4440
ACTCCAGGGGCACTACCTCTTTGTCCAGCCTACTGGGCATAGGCAAGCCATGGAGGTGGA	4500
GGCCATCGGCAACAGGCAGAGGCTGTGTTAGCCAGACATGAAGCTGTGGCTGGCACTC	4560
GGAGCAAACGCTGCTGAAGAGGGCCAAGCGCAGTATCCACTTCAATGATCCCAAGTACCC	4620
TCAACAGTGGCACCTGGTAAGTACAGCCTCAGATAGACAGGGCCTTCTCTAGTGTGCTTT	4680

TGACTGAGGCTTCTCCCACTGCCCTGACTTATACTGAAGGGGGCATGTCTGTGGATTG	4740
GGAATATAGTTTGTCTTAGTCCATACCTTCCCTGTCTGCTGTGAAGAGCAGGGCAGTGC	4800
TGAGGGTGATGCACTCAACTGCAGTAGTCTACTGCACCTCCTGTTCTAGCCTTCCACTT	4860
TCTAGCATGTACCTTTGGCTGAACTATAACCCCTCTGTGCCCACTTGTAATAATGGAGGTT	4920
GTAGCAGTGCATATTTCAATGGAGTACTGTTAAAAACAAGTAAGTGTGTATGTGAAGCATT	4980
TAGAAAAGTGTCTTACTCATGGGGAAACCTGTGTATTGGCTGATTGCAAACTTTAC	5040
TGCATAATAATTACCAATTAGCCCTAGCTACCACTTTGTCTTGAGTTGCCTACAAGTGA	5100
GACCAAGAATCGATCGGAGGGACAGGACTCTCATGTGCCCATCTGGACTCTGTCCCCAC	5160
TAACCTGTTGGTCTGTAGCCTCTGTTGGCTCTTTAATAAAAGCCAAGGAAAGGGGTAGC	5220
TTAACTTGCTCTCTGCACTGAGCTGATGTCCGGTTGATTCTGTCTTCTACAGAAATAA	5280 Exon 4
TCGGCGGAGCCAGGAAGAGACATCAATGTGACAGGTGTGTGGGAGCGAAATGTAACTGG	5340
CCCTGGGGTGCAGTGGTGGTGGATGACGGAGTGGAAACACACCCGTCCAGGACATCGC	5400
ACCCAACCTACGTGAGTGGCTGAAGTGAGCCCTGCTGCCACCCTGCTCCTTCCCTTTTGT	5460
TCCCTTTCCTAGAAAGCACTGAGCCAAAGCAGACAGCCTGGGAGGGGATAGTGTCTGCC	5520
TCCTTTCCTCAGTCAGCTTCCAGACTTGCCTTAAGGCCCGAAGTTGACTTTGGTCTAT	5580
AGCATTAAATCCCTGCATCCTGTTGTGTGCTTTCTGGGAGGGGAGCAGGGAATGCCCC	5640
TATAGGCACTACTTCTCTTTAGAGTTCTGAAAAACCTCTGGCTAAGGTGTGATCTTTA	5700
GAAAGTCTGAAATCTTTACCCCTGGGCTCCCCCAAGCAGCAGGCATAGTACTATAACCAG	5760
CCAAGAGAAAAGCTTGATTAATAAACAAACAAAACAAACACCCCTGCCCTCTCAGCTTCC	5820
CTCCAGTGTCTTGGCTCTTACCTTATAACCTTGGTTGTCTCGAGGCAGTGGGTGTGCTA	5880
CCCTGATTTGGGAGTGGAACTTCCAGTTCCTATTCTCCCTGATGTTAGGAACTGACTCC	5940
TCCTAACCCAGTAAGTAACCTTTGTTGACCCCTCAATGTTTCAAGTCCAGAGGGAAGCT	6000 Exon 5
ATGACCTCAACTCTAATGACCCAGATCCCATGCCCAACCTGATGAGGAGAATGGTAACC	6060
ACCATGGGACCCGGTGTGCAGGAGAAATTGCAGCTGTGCCCAACAACAGCTTCTGTGCAG	6120
TGGGTGTGGCTACGGGAGCCGAATAGCAGGTAACCTCATGCTTCTCCTATCTCTGGGAAG	6180
CTCTTGGGACACAGTGATCATTCTACCTTCCCCATTCTGAAGTTTCGTTTTCATACCC	6240
TCAGTGTTTGATGCTAAAGTGAGGAGTAAGAGCATAGCAAGTTAATTAGCTCATTGACT	6300

TTTCAGGGCTGTGGTATGTCACATGGGCTTTTAACAACAGTATTTATTGTACAGTAGAAA 6360
 ACAGGTAAGTGTGCAAAAGGCTCATATGTGGGCTATAGGTCCCAGGCTGCCCTGCTCAA 6420
 GTCTCTCTTCTGTTTCTGATCATGGCTCTGCTCACCACGACCCACGTCAGTACBCTTC 6480
 CTTGAGTTGACTGTGCTCTGCTGCTGTTTTCATATGTTATCAAGTGCCTCCCTCATAGT 6540
 TTGCCTCCACAACCTCCACATYATGGAGGCTGCTCTCTCAGCCCCACTGCTCCAGTCCC 6600
 CTGCTCTCCATTCTCTTAAGAGAGCTTAGTTGTCTGGCTTTTAAGAGAGGTAGTTATATG 6660
 GTTCYTCATCTCCCAGTCTTGTGGGTCAAG6GGAAAGGATGGGCACTCCTGAATCCTGGG 6720
 AGCCTATGGTCTCTACAGGCTGGCTGTCACCCTGGTCTTTTCCAGCAGGGGGCCTCAGCA 6780
 CTCTGCAAATCCCAATTTGTTCTTAACCTTGACCCACAGTGACCCACTCACTCCATCTTA 6840
 GAAAAAGAGTTTGAGCAGCACTGGGGTAGCAAAAGGGGCGAATGCTATTTACTTCATTTCT 6900
 CTATTGTGTCAGTCACTGAGCTAGAGGTTCCACATGATGTGCCAAGCTGAGRGGGGATAT 6960
 CTTAGATTACCTACCTATATTCTAAACCCAGTGCTTTTCTCACCTCTCAGGATCGTTAGAA 7020
 TCTGAGTTCAAGCATGGTAGTAATAGACACTTGAAAGCTAGGCATATCCCAGAATTTGA 7080
 GAAGTAGGGCCAGGAGGATCAGGGATTCAAAATCATCTTTCTGCTATGTTGTAAGTTAA 7140
 GACAAGACTATGCTGGCTAAGACTCTGTCTTTATAAAATAAAATAAAATAAAATGGCAGGC 7200
 TGTAGTGGCATAATGCTTTTAATCCTAGCACTCAGAAAGCAGAAACAGGCAGATCTCTTGT 7260
 TCGAGGCCAGCCTGGTCTATAACATGAGTTCTAGGAGAACCAGGACTACACAGAGAAACC 7320
 CTGTCCCCAAAAAACCCAAGTAAACAAAAACCATTTTTGAACTAGTAATACTTGTAGAT 7380
 GGTATAAAAAATCCAAAGTACAAAAGAGGGGGTTGGTGTGTAGTTGATAAAGTGCTTGCC 7440
 AGGCAAGCATGAAGACCTGAGTTCAATTCCTGGCACCCATGAAAAAGCTAGAAACAATT 7500
 GGGTTTTCTTGTAATCCCAGAAGTGAACACAGGGATAGATGAATCCATGGGGCTCTCT 7560
 GGATAGCCTGCCTAACTTAATGGGTCCCAAGTCCCATCGAGAGGCTCTGTGTCCAAAG 7620
 ACAAGTATTCTGAAGAATGCTATCCCAAGATAGACGACCACACACACTTACCCGACACAT 7680
 ATGCATATGTACTAATAATGCATTCACACCCCTACATATATAGCTATGAAXRATATGCACA 7740
 CACAXAAGTGCTAAAGAACATACCACCCCTACCACCTGGCACCCCTACTTGGGGAATGTGCA 7800
 CTATTTCTAGTTCCCTTATGGGACCAGCATCAGTGAGGAGCTGGTGTAACCTCAATACCTA 7860
 GTGATCCGGGTGCTGGATGGACCCTTACTGACAGTATGGAGGCTGTGGCGTTCAACAAG 7920 Exon 6

CACTATCAGATCAATGACATCTACAGCTGCAGGTAAGGAGGGGTGAGCTCCAGGGGAGGA 7980
 GGTGAGAGACTCGAGTCTCTGGGACAGATCTTCAAAAAAGAAAAAGAAGCATAGGGA 8040
 AACAGGCTGCCTCTAGCCTAGCAGGGGTTCTTGTTAGACTCAGCTAATGAGGCCATTTT 8100
 TCTTCATTTATCTAGTTCGCCAGCCGCTGGGTGCTGGGTCTTAGACTTGCCACAAAAAGA 8160
 GTGTCCCAGCATGCATGCAATTTGTTGTAAATTCATAATGCTCATAGAAGAAAGTGCAC 8220
 TGAGTTCAGTGCCCCCTGCAAAGATGCACTTGAAGCTTCAGGCTGTTGGTTTAGGAGTT 8280
 GGATAGGAGGAGTTAAGATGGATGGCCCTGTCATATTGATAGAGCCTGTTCTTCCCCT 8340
 AACCAAGTACGTTTCTTAAATCTCTTAAATCTCTGTTATTTGGGGGTGGGGTGGTAT 8400
 TGTGCTATTGTGCCAATCCCATGGTCAGAGGACAAGCCCTTATGTGSSCCCTGGGGTC 8460
 AGTCAGTCACTCGGCTTTCATAGCATAACCTTTACCTCCTGAGGCATCTCACAGGCCCT 8520
 GCTCAATTTAAGAACCAGGAATTTGCTGGTGTGCAGCACACATTTTAAATCTCAGAAG 8580
 GCAGAGGCAGATGGATCTCTAAATTC AAGGACAGCCTAGTCTACATGAGTTCAGGACAG 8640
 CCAGGGCTATACAGAGATACTTGTCTCAGTGACAACAACAACAAAAATAACCTTGGAGCT 8700
 AGAGAGATGGCTTGGTGGCTGAGAGCGCTTGATGCTCTTACAGAGGACCTGGTTTCAGTT 8760
 CCCAGTAACTACATGATGGTTTACATCCTTCTGTCACCTCCAGTTTGAGGGAATCTGAAAC 8820
 CTGGCCCTTAGGGGACTCCACATATGTTGAACTTGATACATATTCAGGCAAAATCTT 8880
 AGACAACATAAAAAATAAACTTTAATTAATAAAAAAATAAAATCCAAAACTGAGCAGCTTT 8940
 TATCAAGCCTAGTAGTAGCTCTGTAATCACAGGATGCCGCTGTCCCTATTCCACTACGT 9000
 ATCTTCTGCTTCCCTGGGACCTGAGGCTAGATAAAGTCATTTTAAATTTGGGACATTTCCA 9060
 AATAAAGAAGAGCCTTAGCATATACTTGCAATTTTCTACTTTGGAAGAAAATGGCCCTTG 9120
 GCACCATGTATCTTTAATCATGGGGTTCCTTGGGAGAAGAGTAAAAGCTGCAGGCAGGT 9180
 GAGTGGCTCTGAAGACAAGGAGATCCTTACTCTACTTCCCTTTGCGAATCTCAACAGC 9240 ^{Exon 7}
 TGGGGCCAGATGATGATGGGAAGACAGTGGATGGTCCCTCATCAGCTCGGAAAAGTAACT 9300
 GAGGAGCATGAAGACTGGGGACCCCTTGGGACCCCTTTACTCAGTTTTAGAAGAGTCAT 9360
 CAATCAAAACCCTTCCCATCCCGCTTCATTATAGTGCCCAACTTACTCCATTTAGCAT 9420
 AGCTGTGCCATTATCCTGTCAAGAAAGATTGTGTGCCTCACTTAGAATGGAGGACTGTA 9480
 GGACCCTGGTCTCAGAGAAGGAATAACTTCCAGCTTGTTCCTTTGGATCTCACTCTGG 9540

GTTCTGTCCCTCACCCCTGCTCAGTTATCTGATAAGCAGAACATGTCTGTCAAAGACTAT 9600
 TGGTTTTTTGGGGGGTTGGGGTTACGTTTTTTAGATTATTGATTTTTATATTCAGTGG 9660
 TATTTTCCCTGCATGTATATCTATGTGAGGGGTGCAAGATCCTCCTGGAAGTAGAGTTA 9720
 CAAGACAGTTGTGAGCTGTCATATGTGTTCTGAGAATTGAATTTGGCCCCCTCTGAAGAAC 9780
 AGCCAGTGTCTTAACCGCTAAGCCATCTCTCCAGCCCAAACACTGTTAAGTCTTAGCTT 9840
 TCTCCCCACTTTATGTATGTGTATAAATCTCTTTCCCCCTTGCCCAAGGCCTCACCCCT 9900
 TGTCTTTCTTTCTAGATGCCCCACCCCTACAATCAATCCTAGATGGTCCAGACACCTCA 9960
 TCCCAGACAGCCACATCATTCCAGACACACTGTAGAAGTTACAGCACTGATAGGCCAGCC 10020
 ATAGTCTTCAAGCAAGATCCTCGAGCTCTAGAAGCAGGAAGAGTGGATGGCTGGTTCAAT 10080
 GGGTACAGGTTTTTCTATTAGCTGCTTGACTAATAGCCTGACATTTTTGTCTTCAGGCT 10140
 GCCTTACAACATGGAGTGTATGGCTGGTCGCCAAGGCTTTGGAAGTATCTTTGTGGTTGCC 10200
 AGTGGTAATGGGGGCTAGCACAATGACAACCTGCAACTATGATGGCTATGCCAACTCCATC 10260
 TACACTGTACCATAGGTAGTGGTCTAGCGATTCCATGGGCATGGGGTACCCGGGTGG 10320
 KTTTGTGTCTCACTGCATTGGCTTCAATCTGTTTTCTCTAAATGTGTAATCAGGATCATCTC 10380
 TTGAGGAACAGTACTAGCCAGCTCTTTTCGAGCTCTCAGGGTTTCTCTGGGACATGTTCA 10440
 ACCATTTCCAGAAATCACCTACCTGTCTTCACTACCATTATGACAGGTTTGAGCTCCAC 10500
 TTTGGTCTCTTTCCATGACCAGGCCACTCTCTTGAAGCTGCCTCTGGCTTCTTTTAAACC 10560
 TGAATTAAGGAATCTGTCTTGCCTCCAGTTATAAAGTTTATCATTTCACCTTTTTATT 10620
 ATTATTTGTTACTCGAGCTCCACATAAATAAGGCTTCTTGACATCAGCCTCATTCTG 10680
 ACTCCTTAGATCAGAAGGAAGTGCCTGCCCTCTCTACCACTGCATACTGTGCCTTC 10740
 CTACCCCTTCCAGGATGTGGCCTTTTCCACATGGCTCTAATATCTCATATGTATGACTCA 10800
 CAGTAGATCTTAGTACACATTGAAGACAAGAGGCTAAGTACAGAGATGCCTGAAGCCTAT 10860
 GGAAGAGTGGTGAATGCAAGAAATCCCTCTGCCATTGAGGCTACCCGCCACATAACA 10920
 GATGAAGCTGTGGTTATTTTTTGTAGCTCATCCCTTCATGATGAAGACAACCTTTGGFATCA 10980
 GGTGCTAGCTATAGGATCCACTTGTATGTGAGCTCTGTCTTGTCTTTTCTGTGCCAGGA 11040
 GCTGTGGATGAGGAGGACGGATGCCTTTTTATGCAGAGCACTGTGCCTCCATGCTGCCA 11100
 GTCACCTTCAGTGGTGGAGACAAGATGCTGCCGAGCATTGTGAGCACTTCTAGGACCCTT 11160

Exon 8

Exon 9

CCCATGTGGTAGTCGGGGTGGTGGTGGGCAGGGGTAKTTCAAGACAAGGACTCACTCTGT 11220
AGCCAGCCAGGCTTGGAACTCATAATATTGTTCAAAGTGGCTTTGAATTTATTAGCCATC 11280
TTCCTGCCTTGGCCTCCATGAGTTGAGATTACAGGTGGGATATACCATACCTAATAAAAAC 11340
AGCTGTTTTTTTACCTCATTCTTTGAGAAAGGGTCTTATAACCTAGGTTGCCTCCAAGT 11400
CACTATGTAGCTGAGGATCCCCTTAGTTCCTGATCCTCCTGCATTAGTGGGGGGTTTTGG 11460
TTTTCTTTTCGAGACAGGGTCTCTCTGTTANNNNNNNNTGTTAGTAAATGTAGAG 11520
TTTGGAGTTAAGCCTGGGTATGGTATTGACCAGGTTTATAATTAAGTTCCTGTCCCA 11580
TTCACAGTTGGCTTCAAGAAGAATTGCTGAATGTTGCTGGGGGAGGAAACATTTTG 11640
TTCTAGTGGCCTTGCTCCTCTGTAATAACCTTCATCCAAGCTCCTGTATTAACTTTA 11700
ATTAACCTTGCTGGTCCCTAAAGAAAAATAAATAAAAAATTAGCTGGGCTGTGGTGGTAA 11760
CATCTTAATCCCAGCACTTGGAAAGCAGAAGCAGGTGGATYTVTDWTBHWAKGCCAGM 11820
SATXGXCCTAXTACAGAGTGAGTCCAGGACAGCCMAGGGCTACACAGAGAAACCCTGT 11880
CTTGAAAAACTAAAAAAGGAAAAAATGGGATGGTCCAGGCCAGAGCTGGCA 11940
GGAGCCAGATGCTTGTGACGATTTGCCAGTGGCCAGTTGGGTTGTTCCCTGAGGTTACTGG 12000
GAATTAAGTAGCTTCATCTTACACAGGGTCAAGTTCATCTTACAAAATGAACCTGTTAT 12060
GCCAGAAAATGACTTCTGGTCAACACATTGTTCTGAGGCTCTGTGCCAGTTGACCCTCAC 12120
CCCAGGGAGGAGAGAGAAGGGATAGACTAAAGTAAAAGCAGAAAAGTTGCCCAAAAATT 12180
GGGGTGATACTAGCCTGTTAGAAACAAGTCCCTATTCTTACCGCCTGGCAGAGTTAT 12240
GTGACCCTCGAATCCTGTCCCTGCAAGGGACCTAACAAAGGCTCTGTAGTAAATGTCA 12300
CTGTTGCTCTACACAGCCTAGCTTTGCCCTTCCAACCTCAACXTGGTTTCTCACCCTCAC 12360
CCCACCTTCCATCTGAACTGTAATTTGTCCAGATTTGTAAGGGAGTTTATTCAGCCAAG 12420
TTCCCATCTTTGGAATGCCCTTGCCCTGGCAACCTCCTGTTGTGTGACCAGGCAGCTGCA 12480
TGGAGCTGTCTGAGGAGAAGGGCCAGGGGAATAGGCTTCCCAGCATCTGGCTTCTTGC 12540
CCAAGGAGCTCCCTAGGTTTTCTTGCCTGTGTGATGTGCTTGCCATTATGGGCAAGT 12600
TGGCAGGCTTTGCATCTGTACTCTAGGCTATCGCTGTGTGAGGCTGTRTCTTAACAGGA 12660
AAAAATCAATCTGTTCCCTGGCACCAGTCAGGCTTCTCTCTGTTGGAGACTACAACGT 12720
GGAGTGTTTTCTTTCTTTTAGGAGAGCATAGTGTCTGGGTGAAAATTACTTCTCTTG 12780

GCTGCTGCAGTAAACATCTGGGATGGGTTGGGGGTGGAGGAAXCAATGGAGAAGGCAGCTC 12840
 TTGGAAGAAGTTGGCACCCTGGTGTCTGACCCAGGAGAGAAGTTGAAAGAAAAA 12900
 GTAGCCTGCAGTGGCTTTATCCACAAAGTTTAGCTGGGCAAGTCCCTGTAAACCCCTTTT 12960
 TGATACACTTATGTTGATGGGGCATGACCTTGATGCCAMACAGATACTGTCTTCTCCA 13020
 GTTTAATGTCTCATCTGAGAGACATGGGTAGGAGAGAGGAGCTCACAGACACCCCRACC 13080
 CAAGACACCACCCAGAGCAGGAAAAGCAGCTTGTCCAGTGGCCAGTATTGCCGGAAG 13140
 ACAGTCCATGTAGATACTGAAACCTGCAGTAGGCCTCAAGGGAAGGTTGTAATAATAA 13200
 TCTTAGGGAACCTGTACCCTACATTCATGGTTCCTCTCCAGCTCCAGCCGGAGCCCA 13260
 GTATAGCTGTAGCCACTCCAGCTAGCTGTGTATGTGAGTGGGATAGAGAGTTAGGA 13320
 ATGAAATGATAAAAAATTAGAGGAGCAGGAACCAAGAACTTATGTGTAACCCACATCCT 13380
 CTCACATCCATTATCTCCTTTCCTTACAGCTGTGCCCCAAGGTCAGTCTCAGTTTCCA 13440
 TCATCTAGAGATGGATGGAAGGGTAAGGGTCTCCTGTGCCTAGGCCTCTGTCATAAGGA 13500
 AAGTTAGAGGAGAGGGTCACCTGAGAGACTTGTTTGAGGACTGTCTCATTCCCTCCAGT 13560
 GTCTAACTCTCCCACTCTTATTCCACCTTGGAGCTGACCACTGACTGGGACCTTCAGA 13620
 AAGGCAGTGGTGCAGTGAAGGCCACACAGGAACCTCAGCTGCAGTCTCTGGCAGCTG 13680
 GAATGATAGCCCTCATGCTGCAGCTGCGGSCCTGGCTCAGATGGCGGATGTACAGCACA 13740
 TAATTGTCTTACAGCCATCCAGGTGAGAGCTGATAACAAGTGCCAGCCACCCCACTC 13800
 CATGGCTGGGCTGTCAGGAGACTCCAGACTCAAAAACCTTCCAAATGTCTCAGGGTGGGA 13860
 TGTGTACAGCATGGGTGTGAGGACAGATCCCAAGATGAGCTGTGGCAAGGCACTCTCAC 13920
 TTGGGAGCTGGTCGTCTGGAGTCCCTCTAGGGAAGCAAGTCACTCATGGAGTTTGACTC 13980
 TTTTTTTTACCTTTTTTGAGGGAATCCATTCTAGCGGCTCAGTGGGAGAAGAAAGTCA 14040
 GGCAATGCTAGAGATTACCAAGGCTTTTATTCTTTTCACTATGAAGATCATATGCAGAC 14100
 TGGCTCACCAATGAGGCTGGCTTCAGCCACAGCCATCAGCATGGTTTGGGCTGCTCAAC 14160
 GCTTGGAGACTTGTCAATGACGCCAAGGTGAACCAAGGTGTAGCCAGCAGGATACCCYG 14220
 TATAGAAAAAAGTGGGGAAGGTACCAATGCCTAAACATCACCTGGCATTTTTATTGGTA 14280
 CTTGACCACITTAGCCCTTATTCTGGCTGATAGCCATCTTATCAGTGGGAAGTCGAAGTT 14340
 CCCAAGTGTGTAGTAGTGAAGGAGAAGTGAATTGGACTCTGGGGCTTTCAGTGTCTTG 14400

Exon 10

Exon 11

TTTTCCTTCCCTTCAGACCATGGCTGTTTCCTCTGAGTCTTAACTTGTFTTTCTCAGGA 14460
 ACATTTGGGGTACAGAAGGGACCTCAAGTGAAGAGTCTGATGATAGAAAAGCTAGAGGG 14520
 TTAGACCTTAGAGAGCTTACTGAGTCTCTGCATGAGTGTGTGTGTGTGTGTGAATA 14580
 CAATAAGTTTTGTTTGTAAATTGCTGTCCCTACTTTTCTTTGAGCAGATCTGGACATCT 14640
 GTTCCTTACTTAGCCTCCTATGTCAGCCCCATGCTGAAGAAAATAAGGCTGTTCSGAGG 14700
 TCCCCCACTCTCTGGAGGTGCTATGGAATGGTAAGCCTCCAGCACAAAGGGTTAATGTC 14760
 AGGTCCAGGCTTGAGGCCGGTGATGGGGAAGCTAAGCTAAGCTCATCCTCCCTGCCAGT 14820
 TCTCCCCCTTCCCTCAGTAGCTTGAGCTGTAGCCAAACCTGTGGAGTTTCTTTGAMCAT 14880
 TTTAGGGCATGTTGCTGCTTCTGARITGGCTGGCCTCTGGCTGCTTAACAAGACCTTTCT 14940
 CCTGAGTTCCTCATATAAGAAGGAGCTGTAAGCCAATGGACTAGGAAGGCAGTGGGCTGA 15000
 GTGGGAGTGGGGCTTAGGGAGCCTCAGCAGGGGTGAAAAAAGTTCTCTGTACTTTTCATG 15060
 CTAAGATCTTGCCAAATACCCTTTCCAGCACAGAGGCCCCAGCATGAAGGAGAGAAATGC 15120
 ATGTTTGGTCTTAAATTTCTCCAGAGAAAGCCAGATGCCATCTCTAAGGGGTGGGGCAC 15180
 AAGCACAAAGCACAGGTTCCCTGGGCHACAGTTTCYACTGTGCTGTTGCTCTTTCTCTCT 15240
 GCTTKCCATCACAGGGTGGGCTGTCTGTGAAGGAGTGCAGGGCACCTTCTGACCCAAGG 15300
 GGTGTGCTCTATGTTAAACCCAGAGGTGGATACCTCCCAAGAGAGACGGGAGGCTGGAT 15360
 TCAGATGTTTTGTGCAGCCTGCATTGTCTTTATTGTTTTGGTGCTTGTGAAAAAACCCAA 15420
 TGCCTCATAACCTAAGCATGCATACTCTACCACTAAGCTACACCCAGCTTCTCTATA 15480
 TGCACATTTCTCAGATGTTCTATCCTTAGTCTCATGGATCAATTTTCAACCTTGAGAA 15540
 ATGTTGACAACCAGAGTTATTTGCTACTTTGTAAAAACTGAGGTGATATGAGCGCCATG 15600
 CTGATATCAGAACTCTGTACTTATTTTCTGTGTCCCTTGAAAGTGCCGAGACCAGAGA 15660
 TATGAGGTCTAGAAGAATGACAAAAGAAGGGACCTTGGGTGGGAGGATAGGCCAGCCTTT 15720
 CTTTATGAACACTTCTCTCCCTAAGCATATCTTACTCCTGTTAGTGCCCTATGTGTGGA 15780
 AGGATACGAGGCCATAGGGAATGCATTTTCCATGAACAGCATGCTGTGCATAGGAAGGGTA 15840
 GAAAACATCCAGGAATGTGGTTTGGGGAAGCCCTTGGCCTTTATACACTGTCTTGGGGAT 15900
 TGTGGGCCTTTGGCCTTACTCAAGGCTGTTTCTCTCACTTCTGTGATGAGCAAG 15960
 TGAACCTCTCAGGGTCAGAGACTTGACCAGCTTTGTTATCCTCTGCTTCCATTATCCAG 16020

Exon 12

GGAAATGTGGCCCTTAAAGTGCCTTCCATAGGTATTGGGCAGCTGGCAACTGCCTGTTG 16080
GTCGCTAGAAAGCATTAAAGAGAAGTGTAGAGGGCAGAATACAGGGGTGGGAGTGCCAAG 16140
CAGGTCATCCATGTCAGAAGCCAGGGAAAGAGGTTTCTTGCCCTTAGTACATTCAGCCCT 16200
CAATAGATAGAACAAAAAGAAAAGGTGAGCCTGAAAGGCTGGCTTGGCCATCAAA 16260
GGAGTGAAATTCACATGGTACAGAAGTGCAGCAATGCAGCCCCAGTCTAGATGCCCCACCC 16320
CCTACAATCAATCCTAGATGGTCCAGACACCTCATCCCAGACAGCCACATCATTCCAGAC 16380
ACACTGTAGAACTTACAGCACTGATAGGCCAGCCATAGTCTTCAAGCAAGATCCTCGAG 16439

上記の配列から解析したマウス LPC 遺伝子のプロモーター領域は次の通りである。

Promoter Region of Mouse LPC

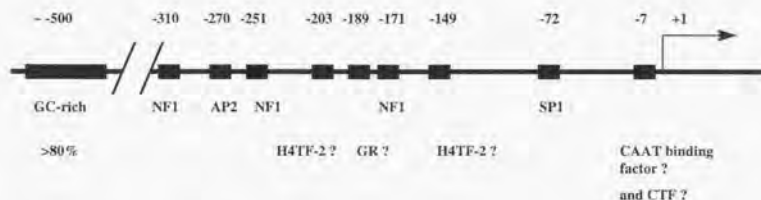


図29 マウス LPC プロモーター領域

プロモーター領域の解析から、SP1 や NF1 といった housekeeping gene に特徴的なモチーフが見つかった。この結果は、Northern blotting の結果とも一致している。また、GR (Glucocorticoid receptor) によって誘導される要素が見つかったが、これが機能しているかどうかは不明である。

<マウス LPC cDNA 配列のシーケンス>

次に、RACE 法によりクローニングし配列を決定したマウス LPC cDNA の配列を示す。

マウス LPC cDNA をシーケンス時には、染色体 DNA をシーケンスした時と同様に、必要に応じて 17~24mer のオリゴヌクレオチドを合成した。

マウス LPC cDNA 配列とアミノ酸配列

1 GC GGC GGC AGC AGC AGC AGC GAC AGT AGC AAC AGA GGC TGT GGG
45 TCT GCA GCT GTG TTC CCA GTC GAC ACT CAC TGC TCA CAG TGA AGG
90 TGA CAT CTC CCT CCA TTT TCA CAA GAT GAG TGT GGT ATG GTT ACC
135 CAG GCC AGG ACC TAA ACT TCA CAG AGA GCC CAG CCT GCT GTT CTG

180 ATG CCG AAA GGG AGG CAG AAA GTC CCA CAC TTG GAT GCC CAC CTG
1 Met Pro Lys Gly Arg Gln Lys Val Pro His Leu Asp Ala His Leu

225 GGC CTG CCC ATC TGC CTC TGG CTG GAA TTA GCC ATC TTC TTT CTG
16 Gly Leu Pro Ile Cys Leu Trp Leu Glu Leu Ala Ile Phe Phe Leu

270 GTT CCC CAG GTC ATG GGC CTA TCA GAG GCA GGT GGG CTT GAC ATC
31 Val Pro Gln Val Met Gly Leu Ser Glu Ala Gly Gly Leu Asp Ile

315 TTG GGC ACA GGG GGG CTG AGC TGG GCC GTA CAT CTG GAC AGC CTA
46 Leu Gly Thr Gly Gly Leu Ser Trp Ala Val His Leu Asp Ser Leu

360 GAA GGT GAG AGG AAG GAA GAG AGT CTG ACG CAA CAG GCC GAT GCT
61 Glu Gly Glu Arg Lys Glu Glu Ser Leu Thr Gln Gln Ala Asp Ala

405 GTG GCC CAG GCA GCA GGG CTT GTG AAT GCT GGG CGC ATT GGA GAA
76 Val Ala Gln Ala Ala Gly Leu Val Asn Ala Gly Arg Ile Gly Glu

450 CTC CAG GGG CAC TAC CTC TTT GTC CAG CCT ACT GGG CAT AGG CAA
91 Leu Gln Gly His Tyr Leu Phe Val Gln Pro Thr Gly His Arg Gln

495 GCC ATG GAG GTG GAG GCC ATG CGG CAA CAG GCA GAG GCT GTG TTA
106 Ala Met Glu Val Glu Ala Met Arg Gln Gln Ala Glu Ala Val Leu

540 GCC AGG CAT GAA GCT GTG CGC TGG CAC TCG GAG CAA ACG CTG CTG
121 Ala Arg His Glu Ala Val Arg Trp His Ser Glu Gln Thr Leu Leu

Cleavage site

585 AAG AGG GCC AAG CGC AGT ATC CAC TTC AAT GAT CCC AAG TAC CCT
136 Lys Arg Ala Lys Arg Ser Ile His Phe Asn Asp Pro Lys Tyr Pro

630 CAA CAG TGG CAC CTG AAT AAT CGG CGG AGC CCA GGA AGA GAC ATC
151 Gln Gln Trp His Leu Asn Asn Arg Arg Ser Pro Gly Arg Asp Ile

675 AAT GTG ACA GGT GTG TGG GAG CGA AAT GTA ACT GGG CGT GGG GTG
166 Asn Val Thr Gly Val Trp Glu Arg Asn Val Thr Gly Arg Gly Val

720 ACA GTG GTG GTG GTG GAT GAC GGA GTG GAA ACA CAC CGT CAG GAC
181 Thr Val Val Val Val Asp Asp Gly Val Glu Thr His Arg Gln Asp

765 ATC GCA CCC AAC TAC AGT CCA GAG GGA AGT TAT GAC CTC AAC TCT
196 Ile Ala Pro Asn Tyr Ser Pro Glu Gly Ser Tyr Asp Leu Asn Ser

810 AAT GAC CCA GAT CCC ATG CCC CAC CCT GAT GAG GAG AAT GGC AAC
211 Asn Asp Pro Asp Pro Met Pro His Pro Asp Glu Glu Asn Gly Asn

855 CAC CAT GGG ACC CGG TGT GCA GGA GAA ATT GCA GCT GTG CCC AAC
226 His His Gly Thr Arg Cys Ala Gly Glu Ile Ala Ala Val Pro Asn

900 AAC AGC TTC TGT GCA GTG GGT GTG GCC TAC GGG AGC CGA ATA GCA
241 Asn Ser Phe Cys Ala Val Gly Val Ala Tyr Gly Ser Arg Ile Ala

945 GGT ATC CGG GTG CTG GAT GGA CCA CTT ACT GAC AGT ATG GAG GCT
256 Gly Ile Arg Val Leu Asp Gly Pro Leu Thr Asp Ser Met Glu Ala

990 GTG GCG TTC AAC AAG CAC TAT CAG ATC AAT GAC ATC TAC AGC TGC
271 Val Ala Phe Asn Lys His Tyr Gln Ile Asn Asp Ile Tyr Ser Cys

1035 AGC TGG GGC CCA GAT GAT GAT GGG AAG ACA GTG GAT GGT CCT CAT
286 Ser Trp Gly Pro Asp Asp Asp Gly Lys Thr Val Asp Gly Pro His

1080 CAG CTC GGA AAG GCT GCC TTA CAA CAT GGA GTG ATG GCT GGT CGC
301 Gln Leu Gly Lys Ala Ala Leu Gln His Gly Val Met Ala Gly Arg

1125 CAA GGC TTT GGA AGT ATC TTT GTG GTT GCC AGT GGT AAT GGG GGC
316 Gln Gly Phe Gly Ser Ile Phe Val Val Ala Ser Gly Asn Gly Gly

1170 CAG CAC AAT GAC AAC TGC AAC TAT GAT GGC TAT GCC AAC TCC ATC
331 Gln His Asn Asp Asn Cys Asn Tyr Asp Gly Tyr Ala Asn Ser Ile

1215 TAC ACT GTC ACC ATA GGA GCT GTG GAT GAG GAG GGA CGG ATG CCT
346 Tyr Thr Val Thr Ile Gly Ala Val Asp Glu Glu Gly Arg Met Pro

1260 TTT TAT GCA GAG GAG TGT GCC TCC ATG CTG GCA GTC ACC TTC AGT
361 Phe Tyr Ala Glu Glu Cys Ala Ser Met Leu Ala Val Thr Phe Ser

1305 GGT GGA GAC AAG ATG CTG CGG AGC ATT GTG ACC ACT GAC TGG GAC
376 Gly Gly Asp Lys Met Leu Arg Ser Ile Val Thr Thr Asp Trp Asp

1350 CTT CAG AAG GGC ACT GGC TGC ACT GAA GGC CAC ACA GGA ACC TCA
391 Leu Gln Lys Gly Thr Gly Cys Thr Glu Gly His Thr Gly Thr Ser

1395 GCT GCA GCT CCT CTG GCA GCT GGA ATG ATA GCC CTC ATG CTG CAG
406 Ala Ala Ala Pro Leu Ala Ala Gly Met Ile Ala Leu Met Leu Gln

1440 GTG CGG CCC TGC CTC ACA TGG CGG GAT GTA CAG CAC ATA ATT GTC
421 Val Arg Pro Cys Leu Thr Trp Arg Asp Val Gln His Ile Ile Val

1485 TTC ACA GCC ATC CAG TAT GAA GAT CAT CAT GCA GAC TGG CTC ACC
436 Phe Thr Ala Ile Gln Tyr Glu Asp His His Ala Asp Trp Leu Thr

1530 AAT GAG GCT GGC TTC AGC CAC AGC CAT CAG CAT GGT TTC GGC CTG
451 Asn Glu Ala Gly Phe Ser His Ser His Gln His Gly Phe Gly Leu

1575 CTC AAC GCT TGG AGA CTT GTC AAT GCA GCC AAG ATC TGG ACA TCT
466 Leu Asn Ala Trp Arg Leu Val Asn Ala Ala Lys Ile Trp Thr Ser

1620 GTC CCT TAC TTA GCC TCC TAT GTC AGC CCC ATG CTG AAA GAA AAT
481 Val Pro Tyr Leu Ala Ser Tyr Val Ser Pro Met Leu Lys Glu Asn

1665 AAG GCT GTT CCG AGG TCC CCC CAC TCT CTG GAG GTG CTA TGG AAT
496 Lys Ala Val Pro Arg Ser Pro His Ser Leu Glu Val Leu Trp Asn

1710 GTC AGC AGG ACG GAC CTG GAG ATG TCG GGG CTG AAG ACC TTG GAA

511 Val Ser Arg Thr Asp Leu Glu Met Ser Gly Leu Lys Thr Leu Glu

1755 CAT GTG GCA GTG ACA GTC TCC ATC ACT CAC CCA CGA CGT GGC AGC

526 His Val Ala Val Thr Val Ser Ile Thr His Pro Arg Arg Gly Ser

1800 TTA GAA CTG AAA CTG TTT TGT CCC AGT GGC ATG ATG TCT TTG ATC

541 Leu Glu Leu Lys Leu Phe Cys Pro Ser Gly Met Met Ser Leu Ile

1845 GGC GCG CCC CGC AGC ATG GAC TCG GAT CCT AAT GGC TTC AAT GCT

556 Gly Ala Pro Arg Ser Met Asp Ser Asp Pro Asn Gly Phe Asn Ala

1890 TGG ACC TTT TCC ACT GTG CGG TGC TGG GGG GAA AGA GCA AGA GGC

571 Trp Thr Phe Ser Thr Val Arg Cys Trp Gly Glu Arg Ala Arg Gly

1935 GTC TAC AGA CTG GTT ATC AGG GAT GTA GGA GAT GAG CCG CTC CAG

586 Val Tyr Arg Leu Val Ile Arg Asp Val Gly Asp Glu Pro Leu Gln

1980 ATG GGC ATC CTC CAG CAG TGG CAG CTG ACC CTG TAT GGC TCC ATG

601 Met Gly Ile Leu Gln Gln Trp Gln Leu Thr Leu Tyr Gly Ser Met

2025 TGG AGT CCA GTA GAC ATC AAG GAC AGA CAA AGT CTC TTA GAA AGT

616 Trp Ser Pro Val Asp Ile Lys Asp Arg Gln Ser Leu Leu Glu Ser

2070 GCT ATG AGT GGA AAA TAC CTG CAT GAT GGC TTC ACC CTG CCT TGC

631 Ala Met Ser Gly Lys Tyr Leu His Asp Gly Phe Thr Leu Pro Cys

2115 CCA CCT GGG CTG AAA ATC CCT GAG GAG GAT GGC TAC ACC ATT ACC

646 Pro Pro Gly Leu Lys Ile Pro Glu Glu Asp Gly Tyr Thr Ile Thr

2160 CCT AAC ACA CTC AAG ACC CTT GTG CTG GTA GGC TGC TTC AGT GTC

661 Pro Asn Thr Leu Lys Thr Leu Val Leu Val Gly Cys Phe Ser Val

2205 TTC TGG ACT ATT TAC TAC ATG CTA GAA GTG TGC CTG AGC CAG AGG

676 Phe Trp Thr Ile Tyr Tyr Met Leu Glu Val Cys Leu Ser Gln Arg

2250 AAC AAG GCT TCC ACC CAT GGC TGC AGG AAG GGA TGC TGC CCC TGG

691 Asn Lys Ala Ser Thr His Gly Cys Arg Lys Gly Cys Cys Pro Trp

2295 GCC CCA CGA AGG CAA AAC TCC AAG GAT GCA GGG ACA GCA CTA GAA

706 Ala Pro Arg Arg Gln Asn Ser Lys Asp Ala Gly Thr Ala Leu Glu

2340 TCA ATG CCA CTG TGC AGC AGC AAG GAC CTG GAT GGA GTG GAT TCA

721 Ser Met Pro Leu Cys Ser Ser Lys Asp Leu Asp Gly Val Asp Ser

2385 GAG CAT GGG GAC TGC ACA ACT GCA TCT AGC TTC CTG GCC CCA GAG

736 Glu His Gly Asp Cys Thr Thr Ala Ser Ser Phe Leu Ala Pro Glu

2430 CTG GAT TGT CCT CCA CAC CAA CCC CCA GAC CTG CTG CAG GGG AAG

751 Leu Asp Cys Pro Pro His Gln Pro Pro Asp Leu Leu Gln Gly Lys

2475 AGC GGG CAG ATT TGC TGA CCC CAG AGC CCA GCT TCC CCC ATG TAC

766 Ser Gly Gln Ile Cys

2520 AAC AGG CAC TTC CTA AAC TCG GTT ATG AGG CTT TCA ATC GTG GAA
2565 GCT CAG GAG GGA ATG GTC CTG ATA GTG ACA TCT GGG ACC CCA GGC
2610 CTC TGA AGC ATT TGC AGC CTC CTC AGG AGG GTG AGG GCC ATC TTA
2655 ATG CAG TGG TAA GGG ACT GGT GAT CAT GGC TTC TCT AGC TGA ACT
2700 TGG CTG AGA ACC TTT GAC AAG GGA TTG GCA CCA AGA CCA GGC AGC
2745 CCT TGA TTT CCT CGC TCT CTG GAG CTA GTC AGT GTC CTA GAC CAC
2790 CAG GAC CTC TAC TCA TTC ATG AGT TAG TTA TGC AGC AAG TCC CTA
2835 TAT ACC TAG AAA TGC TCT TGT CAG AAG TGT CAC TGC CAA GGG ACC
2880 AGA GAG AGG CTG CTA AGA GCC TGG ACA TGC CTA CCC TGA AAG CAG
2925 CTG CCT TCA TTA CCC TTT TCT GTG CGC CTG TTC TGA AGC CAA CTT
2970 AAA CCT GTC CAC ATC TTC ACT ACC ATT CGC CCT TCT GCT GAA GGA
3015 ATG GTG ACG TGT TCC CAG GAG AAC TGG GAG GCA ACA GGA TTC CCC
3060 AGA CCC TTG CTG CCT CAT TGG GAA GTT GGC CAG CTG GAC TTC AGA
3105 GAG CTA AGA CAC AGC ATT TTC AAA GAT AGC CTG GGA GCA CTG TCT
3150 CTG ACC ACC TGT CTT CCT CTG CAC CGT GCT CAG GGA AAT GGC CTT
3195 CAT TCC AGA GGC CAG CTG TCT GCC TGA CTT TTC CTC CGT TCT TGG
3240 CCT TCT CCC CTT CCC ATC TTC TGA GCT AAT TGT TAA TAT GAA TTT
3285 TTT AAT GCT TAA GAT TTG ATT TTT ACT TTT CAA AGC AAC GAT TTT
3330 GTT GAA TTT TTT TCT GCA CAG CTT TCC AAA ATA AAA AGC AGA AGT
3375 ATA AAA ACC CAA AAA A

ポリ A 鎖付加シグナルは、下線で示した。また、proprotein convertase に保存されている活性に関与しているアミノ酸（4アミノ酸残基）は、四角で囲って示した。

Catalytic domain は下線、middle domain には破線で示した。

また、LPC が成熟化する際に切断される部位は縦線で示した。

以下にマウス LPC cDNA とヒト LPC cDNA の塩基配列を示す。

上段がマウス LPC、下段がヒト LPC 配列である。また、マウスおよびヒトで保存されている配列は「|」で示した。一致しない塩基は、空白で示している。

ギャップを挿入することにより、配列の alignment を最適化している。これは、特に 3'末端の共通配列を見つけるために有効である。3'non-coding 領域の共通配列は、mRNA の安定性や他の因子が結合するなど、何等かの役割をもつと考えられる。

AGAGCCCAG-CCTGCTGTCTGTATGCCGAAAGGGAGGGCAGAAAAGTCCCACACTTGGATGCCCACTGGGCCCTGCCCATCT 79

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

AGAATCEAGTECA-CTGCTCTGATGCCGAAGGGAGGGCAGAAAAGTGCCACACTTGGATGCCCACTGGGCCCTGCCCACT 79

GCCTCTGGCTGGAATAGCCA—TCCTCTTTCTGGTTCCCCAGGTCATGGGCCTATCAGAGG-CAGGTGGGCTTGACAT 156

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

GCCTCTGGCTGGAATTAGCCGGGCTCTTTACTGGTTCCTGGGTCATGGGCCTGGCAG-GGACAGGTGGGCCTGA—T 156

—CTTGGGCACAGGGGGG-CTGAGCTGGGCGTACATCTGGACAGCCTAGAAGTGA-GAGGAAGGAAGAGAGTCT-GAC 233

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

GGCCAGGGCACAGGGGGGGC-GAGCTGGGCTGTGCACCTGGAAAGCTGGAAGGTGACGGGG-AGGAAGAGACTCTTGA- 233

GCAACAGGCGGATGC-TGTGCCAGGAGGAGGGCTTGTGAATGCTGGGGCATTGGAGAAGCTCCAGGGGACTACCTC 312

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

GCAGCAGGCGGATGCGCT-TGGCCAGGAGGAGGGCTGTAATGCTGGACGCATCGAGAGCTTCAGGGGACTACCTC 312

TTTGTCCAGCCTACTGGGCATAGGCCAATGGAGGTGGAGGCCATGCGGCAACAGGAGGAGGCTGTGTTAGCCAGGCA 392

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TTTGTCCAGCCTGCTGGGCACAGGGGCCCTGGAGGTGGAGCCATCCGGCAGCAGGTGGAGGCTGTGTTGGCTGGGCA 392

TGAAGCTGTGGCTGGCACTGGAGCAACGGTGTCTGAAGA-GGGCCAAGCCAGTATCCACTTCAATGATCCCAAGTAC 471

|||||
|||||

TGAAGCTGTGGCTGGCACTGAGAGCAGAGGCTGCT-AAGGCGGGCCAAGCCAGCGTCCAATTCAACGACCCCAAGTAC 471

CCTCAACAGTGGCACCTGAATAATCGGCGGAGGCCAGGAAGAGACATCAATGTGACAGGTGTGTGGGAGCGAAATGTAAC 551

|||
|||

CCGCAGCAATGGCACCTGAATAACCGACGGAGCCGGGACAGGACATCAACGTGACGGGTGTGTGGGAACGCAATGTGAC 551

TGGCGTGGGGTGACAGTGGTGGTGGATGACGGAGTGGAAACACACCGTC-AGGACATCCACCCAACTACAGTCCA 630

|||||
|||||

TGGCGCAGGGGTGACGGTGGTGGTGGATGCGGAGTGGAA-CACACCATCCAGGACATTGCACCCAACTATAGCCCT 630

GAGGGAAGTTATGACCTCAACTTAATGACCEAGATCCCATGCCCAACCCCTGATGAGGAGAATGGCAACCACCATGGGAC 710

|||||
|||||

GAGGGTAGCTATGACCTCAACTTAATGACCEAGATCCCATGCCCAACCCCTGATGAGGAGAATGGCAACCACCATGGGAC 710

CCGGTGTGCAGGAGAAATTCAGCTGTGCCCAACAACAGCTTCTGTGCAGTGGGTGTGGCCCTACGGGAGCGGAATAGCAG 790

|||
|||

CCCATGTGCAGGAGATCGGGCTGTGCCCAACAACAGCTTCTGTGCCGTGGGCTACGGGAGCGGCATGGCAG 790

GTATCCGGTGTGGATGGACCCTTACTGACAGTATGGAGGCTGTGGCGTTCAACAAGCACTATCAGATCAATGACATC 870

|||||
|||||

GTATCCGGTACTGGATGGACCCTTCTCAGCAGCAGCATGGAGGAGTGGCGTTCAACAAGCACTATCAGATCAATGACATC 870

TACAGCTGCAGCTGGGGCCAGATGATGATGGGAAGACAGTGGATGGCTCATCAGCTCGGAAAGGCTGCCCTACAACA 950

|||||
|||||

TACAGCTGCAGCTGGGGACCAGATGACGATGGGAAGACAGTGGATGCCCCATCAGCTTCGAAAGGCTGCCCTACAACA 950

TGGAGTGATGGCTGGTCGCCAAGGCTTTGGAAGTATCTTTGTGGTTGCCAGTGGTAATGGGGCCAGCACAATGACAAC 1030

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

TGGGGTGATTGCTGGTCGCCAGGCTTTGGGAGCATCTTTGTGTTAGCCAGTGGCAACGGAGGCCAACACAACGACAAC 1030

GCAACTATGATGGCTATGCCAACTCCATCTACACTGTGCACCATAGGAGCTGTGGATGAGGAGGGACGGATGCCTTTTAT 1110

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

GCAACTACGATGGCTACGCCAACTCCATCTACACCGTCACCATAGGAGCTGTGGATGAGGAGGGACGGATGCCTTTCTAT 1110

GCAGAGGAGTGTCCCTCCATGCTGGCAGTCACCTTCAGTGGTGGAGAGAAGATGCTGCGGAGCATTGTGACCACTGACTG 1190

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

GCAGAAGAATGTGCTCCATGCTGGCAGTCACCTTCAGTGGTGGGACAAGATGCTTCGGAGCATTGTGACCACTGACTG 1190

GGACCTTCAGAGGGCACTGGCTGCATGAAGGCCACACAGGAACCTCAGCTGCAGCTCCTCTGGCAGCTGGAATGATAG 1270

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

GGACCTTCAGAGGGCACTGGCTGCATGAGGCCACACAGGACCTCAGCTGCAGCGCCTCTGGCAGCTGGCATGATAG 1270

CCCTCATGTGTCAGGTCGGGCCCTGCCTCACATGGCGGATGTACAGCACATAATGTCTTCACAGCCATCCAGTATGAA 1350

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

ECTTAATGTCTCAGGTCGGGCCCTGCCTCACGTGGCGTGACGTGCAGCACATCATTTGTCTTCACAGCCACCCGGTATGAG 1350

GATCATCATGCAGACTGGCTCACCAATGAGGCTGGCTTCAGCCACAGCCATCAGCATGGTTTCGGCCTGCTCAACGCTTG 1430

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

GATCGCGTGCAGAGTGGTCACCAACGAGGCGAGGCTTCAGCCATAGCCACCAGCAAGGTTTCGGCCTCCTCAACGCTTG 1430

GAGACTGTCAATGCAGCCAAGATCTGGACATCTGTCCCTTACTTAGCCTCCTAATGCAGCCCATGCTGAAAGAAAATA 1510

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

GAGGCTCGTGAATGCAGCCAAGATCTGGACATCTGTCCCTTACTTAGCATCTACGTGAGTCCCGGTTAAAGAAAACA 1510

AGGCTG-TTCCG-AGGTCCCCTACTCTCTGGAGGTGCTATGGAATGTGACGAGGACGGACCTGGAGATGTGGGGCTGA 1588

|||| | ||||| || |||||| | | ||||||| || ||||||||||||| ||||||||||||||| |||||||
AGGC-GATTCGGCAG-TCCCCCGTTCCCTGGAGGTCTGTGGAATGTGACGAGGATGGACCTGGAGATGTGAGGGCTGA 1588

AGACCTGGAAACATGTGGCAGTGACAGTCTCCATCACTCACCCACGAGTGGCAGCTTAGAACTGAAAACGTGTTTGTCCC 1668

|||| ||| ||||||||||||||||||||||||||| || | ||| || | |||| || |||
AGACCTGGAGCATGTGGCAGTGACAGTCTCCATCACTCACCCACGCGCGGAGCTTGGAGCTGAAGCTGTTCTGCCCC 1668

AGTGGCATGATGTCTTTGATCGGCGGCCCCGACGATGGACTCGGATCCTAATGGCTTCAATG-CTTGGACCTTTTCCA 1747

||||||| |||| | ||||||| ||||||||||||||||| || |||||||| | ||||||| ||||
AGTGGCATGATGTCCCTCATCGGCGCCCCCGCAGCATGGACTCGGATCCCAACGGCTTCAATGACT-GGACCTTCTCCA 1747

CTGTCCGGTGCTGGGGGAAAGAGCAAGAGCGCTC-TACAGACTGGTTATCAGGGATGTAGGAGTAGCCGCTCCAGAT 1826

||||| ||||||||| |||| | ||| | | |||| | | |||||||| | ||||| | |||| |
CTGTCCGATGCTGGGGGAGAGAGCCGAGG-GACCTACAGGCTTGTATCAGGGATGTGCGGGATGAGTCAITCCAGGT 1826

GGGCATCTCCAGCAGTGGCAGCTGACCCGTATGGCTCCATGTGGAGTCCAGTAGACATCAAGGACAGACAAGTCTCT 1906

||||||| || ||||||||| ||||| ||||| ||||||||| ||||||||| ||||
CGGCATCTCCCGCAATGGCAGCTGACCCATATGGCTCTGTGGAGTGCACTAGACATCAGGGACAGACAAGGCTGT 1906

TAGAAAGTGCTATGAGTGGAAAAATACCTGCATGATGGCTTACCCCTGCCCTGGCCACCTGGGGTGA AAAATCCCTGARGAG 1986

|||| |||| | ||||||||| ||||| ||||| ||||| ||||||||| |||||||
TAGAGAGTCCCATGAGTGGAAAAATACCTGCATGACTTGCCTGCCCTGCCACCGGGGTGA AAAATCCCTGARGAA 1986

GATGGCTACACCAATACCCCTAACACACTCAAGACCCCTTGTGCTGGTAGGCTTCAGTGTCTTCTGGACTATTFACTA 2066

|||| | ||||| |||| | |||| | ||||||||| ||||||||| |||| | ||||||||| |||||||
GATGGTTACACCATACCCCTAACACCCCTCAAGACCCCTGGTGGTAGGCTGTTACCGTCTTCTGGACTATTFACTA 2066

CATGCTAGAAGTGTGCTTGAGCCAGAGGAA—C—AA—G—GC—T—T—CCAC—CC—AT— 2146

||||| ||||| | ||||||||| | || | | | | | |||| | ||

CATGCTGGAAGTATATTTGAGCCAGAGGAATGTGGCTTCCAATCAAGTTTGTAGGAGTGGACCTGCCACTGGCCCATC 2146

GG—CT—GC—AGGAAGGGATGCTGCCCTGGGCCCCAGGAAGGCAAAACCTCAAGGATGCAGGGACAGCACT 2180

|| | || ||||||| | | | | | |

GGAGCCGGAAGCCAAAGGAGGAAGGA—C—A—G—AG—CT 2180

AGAATCAATGCCACTGTGCAGCAGCAAGGA—CCTGGATGGAGTGGATTCAGAG—CATGGGGAG—TGCACAATGCATCT 2252

||||| ||||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

AGAATCAGTGCACCTTTGCAGCAGCAAGGATCCAG—ACGAAGTGGAAACAGAGAGCA—GGGGCCCTCC—CA—C—CA—C— 2252

AGCTTCCTGGCCAGAGCTGGATTGTCTCCACACCAACCCCCAGACCTGCTGCAGGGGAGAGCGGGCAG—ATTTC— 2299

||||| ||||| || || || ||||||||| || ||||| ||||| ||

—CT—C—TGACCTC—CT—TG—CCCCAGACTGTGGAGC—AAG—G—GGACTGGA—GCC 2299

TGACCC—CA—GAGC—CCA—G—CTCCCCCATGTA—CAAC—AGGC—ACTTC—CTAAACTCGGTATGAGGCTTTCA 2369

||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

TGTCCCAAGAACAGAGCCCTGGACTGCCCTCATC—AGCACCTAGACGTAC—CGC—A—CG—GGA—AGGAGGAG—CA 2369

—ATCGTG—GAAGCTCAGGAG—GA—ATGGTCTG—ATAGTGA—CATCT—GGGA—C—C—C— 2442

||||| || ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

GATC—TGCTGAC—CTCAGG—GCCTGACA—G—TG—TGGGACAG—GCTCTCTTCCCAAATTAGGGAGCTCTTGACAGAA 2442

—CAG—G—C—C—TCTG—AA—G—CAT—T—TGCAGCCTCC—TCAG—GAGGGTGAGGGCCATCTTAATG— 2518

||| | | ||||| || | ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

AGCAGTCTGATGCTTACATCTGGAATCTGAGGATCCTCTG—A—C—TCCACTCAAAGAGGGTGAGGGCCTTCTTAA—GA 2518

—CAG-TGGTAAGGACTGGTGATCATGGCTTCTCTAGCTGAACT-TGGCTGAGAACCTTTGACAAGGGAT-TGGCAGCA 2578

| | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TACAAATGGT—GG-A—GG—AT—TG-CTGC-C-AGA-GAAGTCTGG-TCAGAGCC—ACA-GGG-TCTGCCTCCA 2578

AGACCA—GGCAGCCCTTGAT—T-TCCTGCCTC—TCTGGAGC—TAGTCAGTGT—CCTAGACCA 2647

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

—G-CCAAACGGGAGCTTTTGGTGAGAAGGTGT—TGGG-CAGGGGAT-TGGCGCCCCCTT-TG-GTTTGGCCT—CCA 2647

CCAGGAC-C-TCTACTCATI—C-ATG-AGTTAGT-IATG-CAGC-AAG—T—CCCTATATACCT-AGA-AAT 2717

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TC—CTCATCT-CTG-TTGGGCCAAGCCAGCT-GCTAGGTCCCCGAAGCATGGGGACCCC-T-TCCCACATATAA- 2717

GCTCTGTGTCAGAAGTGT-CACTGCCA-AGGGACCAGAGAGAGGC-TGCT-AA-GAG——— 2783

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

G-T-T-G—AGAAG-GTGC-CTGCCATAG—CCAG-GAGGG-CAT-CTCAATG-GAAACATCACTGGGGTCACTTGGGA 2783

————— CCTGGACATGCTACCCCTGAAAGCAGCTGCCTTCATTACCCTTTTCT- 2863

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

AGAGGACTTCGGGGTAGAGGCTGGGAGGAGCCCTGGACATGCCTGTCTGAAAGCGGCTGCCTCCATTATCCATTCCCA 2863

—G-TGGCCCTGTCTGAAGCCAECTTA—AACCTGT-CCACAT—CTTCACT—ACCATTGGCCCTTCTGCTGAA-GGAA 2928

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

AGATGC-C-TGATCAGAAACCAACC-ATGAA—TGAACCCC-TGGCT-C-CTTACCA—C-CCC—CA-C-GATTGGTA 2928

TGGTGACGTGTTCCAGGAGA-A-CTGGGA———GGCA—AC-AGGATCCCCAGACCTTGTGC-CTCATTGGGA 2991

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TGATG-C-TG—CC—GGC-ACAGCTGGGATACACACGGCTCCCCAGG—CCTGAG-C-T-GCTTCACT-A—GGGA 2991

AGT—TGGCCAGCTGGACTTCAGAGAGCTAAGA—CACA—GCAT—T—TTCAA—AGATAGCCT—GGGAGC—ACTG— 3064

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

A—TCCTG—CG—GCAGGACTGCAGAG—C—A—GATGGCAGATGCACATGTTGAGGAGAGAGCCTTGGGAGCCACTGCCAC 3064

TC—TC—TG—ACCA—CCTGTCTTCCTCTGCACC;TGCTCAGGGAAATGGCCTTCATTCCAGAGGCCAGCTGTCTGCCT 3143

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TCCAGTCTGCACCACCCCTGCTTCCTCTGCAA—GTGCTCAGGGAAATGGCCTTCCCGCCGGAGGCCAGCTATCTGCCT 3143

GAC—T—T—T—T—C—CTC—CGTCTTGGCCTTCTGCCCTTCCGATCTCTGAGCTAATGTTAATATGAA—TTT 3210

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

GACAGGCTGTGACTCTTCTCTCAAC—CTTGGCCTTCTCCTCT—CTCTGAGCTAGTTGGT—T—GAATTTTT 3210

TTTAATGCTTAAGATTGATTIT—TACTTTTCAAAGCAAGATTTTGTAATTTTTTCTGCACAGCTTT—CCAAAATA 3288

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

TTTAATGCTTAAGATTGTTTTCT—GTTTTACAGCAAC—ATTJTCTGAAJTTTTTCTGCACAGCTTTTCCAAAATA 3288

AAAAGC—AGA—AGTATAAAAACCCAAAAAA 3321

| | | | | | | | | | | | | | | | | | | |

AAAACCTTCCAAACAAAAAAAATAAAAAA 3321

LPC 全体では、塩基レベルで、ヒトとマウスで 67%保存されており、coding 領域だけで比較すると、83%保存されている。アミノ酸レベルでは、90%保存されていた。この数値は、ほぼヒトとマウスの種の違いを反映した差と考えられる。

また、3'non-coding 領域を比較したところ、ヒト・マウス間のいくつかの領域で高い相同性を示した。しかし、NCBI に登録されているデータを検索してところ、何等かの機能的を示唆するデータは得られなかった。

<マウス LPC3'領域のシーケンス>

マウス LPC3'領域 (エクソン 13~17) のシーケンスは、エクソンとその近傍領域のみ行った。シーケンスに用いた DNA は、マウス LPC のエクソン 13 以降をプローブとしてマウス染色体ライブラリーをスクリーニングして得たファージ DNA およびマウス染色体 DNA そのものである。これらの DNA を用いて、PCR によりイントロンを含む領域を増幅してシーケンスした。

ただし、イントロン 12 (エクソン 12 とエクソン 13 間) は、他のイントロンと比べて大きく、2 回のマウス染色体ライブラリーのスクリーニングにも関わらず、一部しかクローニングすることができなかった。また、エクソン 12 とエクソン 13 の配列から作製したプライマーを用いて long range PCR を行ってもクリアなバンドを得ることはできなかった。この結果は、ヒトのイントロン 12 とまったく同じであった。

表25 マウス LPC のエクソン・イントロン構造

Exon	3'intron~5'exon	3'exon~5'intron	Length of intron
1	GGTGGGGCT-GCGGCGGCAGC	GAGGCTGTGG-GTGAGTCGGGA	1542
2	GGGTCCACAG-GTCTGCAGCTG	GAGAGCCCAG-GTAGGAATGGC	609
3	TTGTTGCAG-CCTGCTGTTCT	GTGGCACCTG-GTAAGTACAGC	639
4	TCTTCTACAG-AATAATCGGCG	ACCCAACTAC-GTGAGTGGCC	574
5	AATGTTCAG-AGTCCAGAGGG	CGAATAGCAG-GTAACTCATGC	1711
6	CAATACCTAG-GTATCCGGGTG	ACAGCTGCAG-GTAAGGAGGGG	1287
7	ATCTCAACAG-CTGGGCCCAG	GCTCGGAAAG-GTAACTGAGGA	843
8	TGTTCTCAG-GCTGCCTTACA	GTCACCATAG-GTAGTGGTCTA	762
9	TCTGTCCAG-GAGCTGTGGAT	GCGGAGCATT-GTGAGCACTTC	3000
10	CACCTTGGAG-GTGACCACTGA	AGCCATCCAG-GTGAGAGCTGA	350
11	ATTCTTCAG-TATGAAGATCA	TGCAGCCAAG-GTGAACCAAGG	400
12	CTTGAGCAG-ATCTGGACATC	CTATGGAATG-GTAAGCCTCCA	?
13	CTTCCTATAG-TCAGCAGGACG	GCATGGACTC-GTATGTTACAC	800
14	ND	AGGGATGTAG-GTGAGCCTCCT	850
15	ATTGCCCTAG-AGATGAGCCGG	ACAGACAAAG-GTAGGTACATC	200
16	TGTCCTGTAG-TCTCTTAGAAA	ND	300
17	ND	ND	—

表中の下線は、スプライシング領域に 100%保存されている配列を示す。
また、上記の表を模式的に表すと、以下の図のようになる。

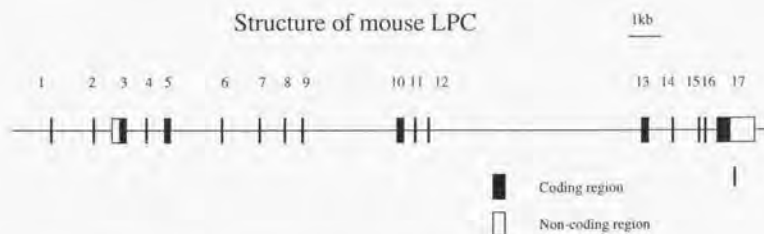


図31 マウス LPC 遺伝子のエクソン・イントロン構造

マウス LPC は、平均して 200bp のエクソンと 1.5kbp のイントロンを持っている（イントロン 12 を除く）。ただし、ORF が始まるエクソン 3 と ORF が終了する 17 が他に比べて大きく、これを除くと、エクソンの平均の長さが 110bp となる。また、エクソン、イントロンを合わせた染色体上の LPC 遺伝子の大きさは、約 28kbp である。

Alternative splicing が起こっているかどうかは明らかにならなかったが、別のグループによると、より短い転写物があるとの報告もある。

4-3-5 マウス cDNA のスクリーニング

ICRF には DNA ライブラリーサービスがあり、必要なライブラリーを提供している。マウス LPC cDNA クローンをスクリーニングするため、ICRF の DNA ライブラリーサービス部より、マウス cDNA クローンが固定されたスクリーニング用メンブレンフィルターを提供してもらい、それをスクリーニングした。スクリーニングには、染色体 DNA をスクリーニングするときに用いたヒト LPC の pro 領域を PCR により増幅した DNA をプローブとして用いた。スクリーニングの結果、合計 4 クローンが得られた。しかし、この 4 クローンは、どれも 1kbp と短く、予想されたマウス LPC cDNA 完全長 3.5kbp と比べてはるかに短かった。また、これら 4 クローンを DNA シーケンスしたところ、「3-3-3 DNA シーケンス」の結果で示されたマウス LPC cDNA 配列とまったく同じ配列が見られたが、同時にどのクローンにも明らかにイントロンと思われる配列も見つかった。また、含まれていたエクソン領域も極めてわずかな領域のみであり、かつ他にマウス cDNA ライブラリーが入手できなかったため、cDNA ライブラリーのスクリーニングをあきらめた。同時に進めていた染色体 DNA のクローニングおよびシーケンスの結果から、マウス LPC 遺伝子の配列の一部が明らかになっていたので、この配列情報を元にして、RACE (Rapid amplification of cDNA

ends) 法によるマウス LPC cDNA の単離を行うことにした。

以下に RACE 法の概略を示す。

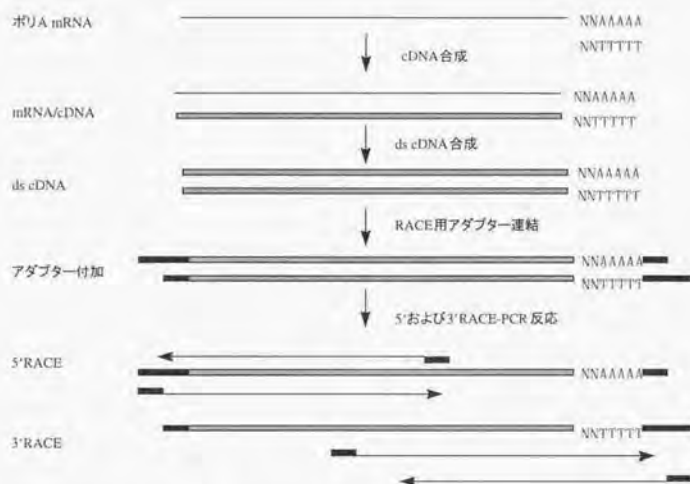


図32 5'および3'RACE反応の概略

RACE 法は、cDNA の一部がクローニングされた時に、残り 5'あるいは 3'領域をクローニングするために使われる方法である。5'領域を増幅する場合、すでにシーケンスされた配列を元に、5'方向へ合成が進むようにプライマーを設計する。また、鋳型となる cDNA には、従来オリゴ dC などが使われていたが、反応ステップ数が多いうえに、付加される C の数をコントロールすることが難しかった。

ここでは、合成した cDNA 末端にアダプターを連結した。これにより、先に合成した 5'方向へ合成が進むようにプライマーとアダプターに相補的なプライマーを使用して目的の遺伝子産物だけが RACE 反応により増幅された。

なプライマーの T_m 値より高い温度をアニーリングおよび伸長反応の温度に用いる。この反応を数サイクル行うことで、目的の cDNA だけが蓄積する。その後、アニーリングおよび伸長反応の温度を下げて反応を行い、増幅産物を得るのである。タッチダウン PCR を行うためには、目的の遺伝子に特異的なプライマー（ここでは MA5 と MA3）の T_m 値が高くなければならない。今回用いた遺伝子特異的プライマー MA5 と MA3 の T_m 値は、ともに 70°C 以上である。

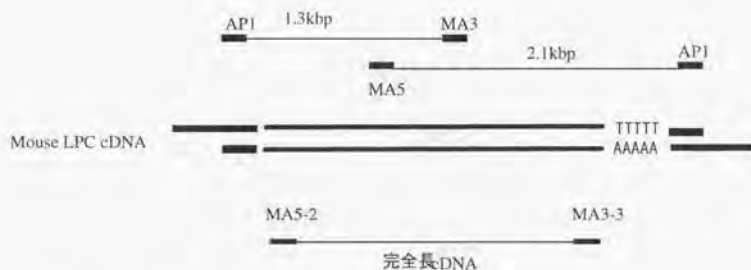


図34 RACE 法とプライマー

図に示したように、すでに既知であった配列を元に MA5 と MA3 を合成した。これらのプライマーとアダプター特異的なプライマー AP1 を用いて RACE 反応を行った（下図の 3'RACE および 5'RACE）。増幅産物を TA-cloning 法によりプラスミドベクターにクローニングした後、DNA シーケンスを行い、cDNA の 5' と 3' 末端に PCR プライマーを合成した（MA5-2 と MA3-3）。当初、完全長の cDNA を得るために 5' および 3' 末端から 30 スクレオチドをプライマーとして用いたが、非特異的なバンドが多く見られ目的の LPC cDNA が増幅されなかった。3' 側のプライマー（MA3-2）の T_m 値が低かったため、アニーリング温度を上げられず、非特異的なバンドを消すことはできなかったと考えられる。

そこで、完全長の cDNA を得ることをあきらめ、十分に高い T_m 値を与えるプライマーが設計できる領域のうち、もっとも 3'末端に近い領域に新しいプライマー (MA3-3) を設計することにした。

この新しいプライマーと MA5-2 を用いることにより、下図のようにほぼ非特異的なバンドを除去することができた (Full length)。このバンドは、Northern ハイブリダイゼーションで得られたバンドよりも小さかったが、ポリ A 鎖および 3'末端から 200bp ほどが含まれていないことから、目的のクローンであると考えられた。



図35 RACE 反応産物

ここで得られた完全長の cDNA は、TA-cloning により、プラスミドベクターにクローニングされた後、塩基配列を決定した (DNA シーケンス参照)。

5 考察

5-1 LPC 遺伝子とその構造

LPC 遺伝子は、ヒトでは 11q23 にマップされる。この領域は、白血病細胞で頻繁に転座が見られる領域である。また、他の腫瘍においても転座が報告されている。このことは、転座の「ホットスポット」であるか、あるいは 11q23 の転座が腫瘍の形成に何等かの形で関与していることを示唆している。

ヒト LPC 遺伝子とマウス LPC 遺伝子の染色体構造を解析した結果、エクソン・イントロン構造はもとより、イントロンの長さまでほぼ同じであることが判明した（下図参照）。

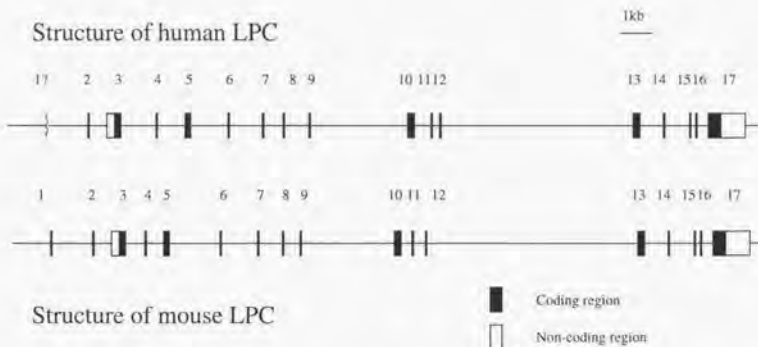


図36 マウスおよびヒト LPC 構造の比較

しかし、他の subtilisin-like protease の中で既にエクソン・イントロン構造が決定されている furin⁶³や PC4⁶⁴とはかなり構造が違うことが明らかになった。furin は、エクソン 1 が複数あると共に、プロモーターのも 2 つあることが報告

されているが、LPC には、alternative splicing 産物の存在や複数のプロモーターの存在を示唆する証拠が見つかっていない。また、PC4 では、prepro 領域がエクソン 1 から 3 までにまたがっているのに対して、LPC では ATG initiation codon がエクソン 3 に存在し、prepro 領域すべてがこのエクソン 3 内に存在する。他の subtilisin-like protease との比較は次節にて行う。

以下に、これまでクローニングされた LPC 遺伝子のアミノ酸レベルの比較を次のページに示す。ただし、

LPC : human LPC

MLPC : mouse LPC

RLPC : rat LPC

XENOLPC : Xenopus LPC

である。

また、四角で囲んだ領域は、subtilisin-like protease の酵素活性中心であり、種を越えて保存されている領域である。Xenopus は、ヒト・マウス・ラットの配列とかなり異なる部分があるが、活性中心部分は、極めて保存されていることがわかる。後述するが、この領域は、他の subtilisin-like protease 間でも、極めてよく保存されている。

LPC のアミノ酸配列全体を比較すると、ヒトとマウス・ラット間で約 88%、マウスとラット間では 96% 保存されているのに対して、ヒトと Xenopus 間では 71%、マウス・ラットと Xenopus 間では 67% が保存されていた。Xenopus およびラットのエクソン・イントロン構造については報告がないので比較はできないが、マウスとヒトとの類似性が明らかになったため、ラット LPC 遺伝子の構造は類似しているものと考えられる。

1

60

LPC. PRO MPRGRQKVPHLDAPLGLPTCLWLELAGLFLLVPMVGLAGTGGPDGGTGG. ASWAVILE
 MLPC. PRO MPRGRQKVPHLDAHLGLPICLWLELA. 1FFLVPQVMGLSEAGGLDILGTGG. LSWAVHLD
 RLPC. PRO MPRGRQKVPRLDARLGLPTCLCLELA. 1FFLVPQVMGLTEAGGLDTLGAGG. LSWAVHLD
 XENOLPC. PRO M. YHIPS. SLCLCCYITFI VSHSPPTHAACSLHNMEDVGSDEL.TWAVHLD

61

120

LPC. PRO SLEGDGEBEETLEQQADALAAQAAGLVNAGRIGELQGHYLFVQPAGHRPALEVEPIRQQVEA
 MLPC. PRO SLEGERKEESLTQQADAVAQAAGLVNAGRIGELQGHYLFVQPTGHRQAMEVEAMRQQAEEA
 RLPC. PRO SLEGERKEESLIQQANAVAQAAGLVNAGRIGELQGHYLFVQPGHGCAMEAEAMRQQAEEA
 XENOLPC. PRO SSACEKGEE. MDTLADELSKATGLENRGVI GELKDHYLF. . . AYRGSTPSQEDARRKAHA

121

180

LPC. PRO VLAGHEAVRWHSEQRLRRAKRSVHFNDPKYPQQWHLNRRSPGRDINVTGVWERNVTGR
 MLPC. PRO VLARHEAVRWHSEQTLLKRAKRS IHFNDPKYPQQWHLNRRSPGRDINVTGVWERNVTGR
 RLPC. PRO VLAKHEAVRWHSEQRLKRAKRS IHFNDPKYPQQWHLNRRSPGRDINVTGVWERNVTGR
 XENOLPC. PRO LFTQHESLTWYSEQKPLKRSKRSLHFNDPKYPQQWHLHNARNPGMDINVTGVWERNVTGR

181

240

LPC. PRO GVTVVVDDGVEITIQDIAPNYSPEGSYDLNSNDPDPMPHPDVENGNHHGTRCAGEIAAV
 MLPC. PRO GVTVVVDDGVIETHRQDIAPNYSPEGSYDLNSNDPDPMPHPDEENGNHHGTRCAGEIAAV
 RLPC. PRO GVTVVVDDGVEITVQDIAPNYSPEGSYDLNSNDPDPMPHPDEENGNHHGTRCAGEIAAV
 XENOLPC. PRO GVTVVVDDGVQHTIQDIQPNYSPEGSYDLNSNDPDPMPHPDGGSDNRHGTTCAGEIAAV

241

300

LPC. PRO PNNSECAVGVAYGSRIAGIRVLDGPLTDSMEAVAFNKHYQINDIYSCSWGPDGDDGKTVDG
MLPC. PRO PNNSECAVGVAYGSRIAGIRVLDGPLTDSMEAVAFNKHYQINDIYSCSWGPDGDDGKTVDG
RLPC. PRO PNNSECAVGVAYGSRIAGIRVLDGPLTDSMEAVAFNKHYQINDIYSCSWGPDGDDGKTVDG
XENOLPC. PRO SNNSECAVGVAYGSRIAGIRVLDGPLTDSMEAFNKHYQINDIYSCSWGPDGDDGKTVDG

301

360

LPC. PRO PHQLGKAALQHGVIAGRQGFSGIFVVASGNGGQHNDNCNYDGYANSIYTVTIGAVDEEGR
MLPC. PRO PHQLGKAALQHGVMAGRQGFSGIFVVASGNGGQHNDNCNYDGYANSIYTVTIGAVDEEGR
RLPC. PRO PHQLGKAALQHGVMAGRQGFSGIFVVASGNGGQHNDNCNYDGYANSIYTVTIGAVDEEGR
XENOLPC. PRO PHQLGKAALQHGVIAGRQGFSGIFVVASGNGGQHNDNCNYDGYANSIYTVTIGAVNEVGR

361

420

LPC. PRO MPFYAEECASMLAVTFSSGGDKMLRSIVTTDWDLQKGTGCTEGHTGTSAAAPLAAGMIALM
MLPC. PRO MPFYAEECASMLAVTFSSGGDKMLRSIVTTDWDLQKGTGCTEGHTGTSAAAPLAAGMIALM
RLPC. PRO MPFYAEECASMLAVTFSSGGDKMLRSIVTTDWDLQKGTGCTEGHTGTSAAAPLAAGMIALM
XENOLPC. PRO MPFYAEECASMLAVTFSSGGDKMLRSIVTSDWNLQKGTGCTEGHTGTSAAAPLAAGMIALM

421

480

LPC. PRO LQVRPCLTWRDVQHIIVFTATRYEDRRAEWVTNEAGFSHSHQHGFLLNAWRLVNAAKIW
MLPC. PRO LQVRPCLTWRDVQHIIVFTAQYEDHFDWLTNEAGFSHSHQHGFLLNAWRLVNAAKIW
RLPC. PRO LQVRPCLTWRDVQHIIVFTATQYEDHRADWLTNEAGFSHSHQHGFLLNAWRLVNAAKIW
XENOLPC. PRO LQVRPCLTWRDVQHIIVFTATRYEDRHAAWETNGAGFSHSHQHGFLLNAWRLVNAAKIW

481

540

LPC. PRO TSVPYLASYVSPVLKENKATPQSPRSLEVLWVNSRMDLEMSGKLTLEHVAVTVSITHPRR
 MLPC. PRO TSVPYLASYVSPMLKENKAVPRSPHSLEVLWVNSRTDLEMSGKLTLEHVAVTVSITHPRR
 RLPC. PRO TSVPYLASYVSPMLKENKAVPRSPHSLEVLWVNSRTDLEMSGKLTLEHVAVTVSITHPRR
 XENOLPC. PRO ESVPYLASYISPVFKENKQIPLVRNTLEVNWVNTAADLRLSGMKTLEHVAVTVTIAHPRR

541

600

LPC. PRO GSLELKLFCPSGMMSLIGAPRSMDSDPNGFNDWTFSTVRCWGERARGTYRLVIRDVGDES
 MLPC. PRO GSLELKLFCPSGMMSLIGAPRSMDSDPNGFNAWTFSTVRCWGERARGVYRLVIRDVGDEP
 RLPC. PRO GSLELKLFCPSGMMSLIGAPRSMDSDPNGFNDWTFSTVRCWGERARGVYRLVIRDVGDEP
 XENOLPC. PRO GNLEIRLSCPSGMSSLJAASRRSDMDPNGFSEWTFSTVRCWGESAEGETYRMVVDTDGDES

601

660

LPC. PRO FQVGII.RQWQL.TLYGSVWSAVDI.RDRQRLLLESAMSGKYLHDDFALPCPPGLKIPEEDGYT
 MLPC. PRO LQMGILLQQWQL.TLYGSMWSPVDIKDRQSLLESAMSGKYLHDGFTLPCPPGLKIPEEDGYT
 RLPC. PRO LQVGILLQQWQL.TLYGSTWSPVDIKDRQSLLESAMSGKYLHDDFTLPCPPGLKIPEEDGYS
 XENOLPC. PRO L.RPGIILQQWQL.TLYGSSWASEDAERRRVEDAQSGKFLNSSFSLPCPPGLVIAEQEGYT

661

720

LPC. PRO ITPNTLKTLLVLCFVFWTVYYMLEVCLSQRNVAASNQV., CRSGCPCHWP., HRSRKAKE
 MLPC. PRO ITPNTLKTLLVLCFVFWTVYYMLEVCLSQRNKASTHG., CRKGCCPWA., PRRQNSKD
 RLPC. PRO ITPNTLKTLLVLCFVFWTVYYMLEVCLSQRSKASTHG., CRRGCCPWP., PQSQNSKE
 XENOLPC. PRO ITNNTLRTLLLGCFAVFWSLYTYLELYWSQSSDGDLDLNCEGINCSWSCSKGTTPSPQE

721

780

LPC. PRO EGTELESVPLCSSKDPDEVETESRGPPTSDLLAPDLLEQGDWSLSQN. KSALDCP. HQH
MLPC. PRO AGTALESMP LCSSKDL DGV DSEHGDC TTASSFLAPEL. DCP PHQP
RLPC. PRO VGTALESMP LCSSKDL DGV DSEHGDC TTASSLLAPELLGEADWSLSQNSKSDLDCPPHQ
XENOLPC. PRO LEAGSESVPLFGVEQEEIEV. TSPLM. YHTSMP SRTQSP

781

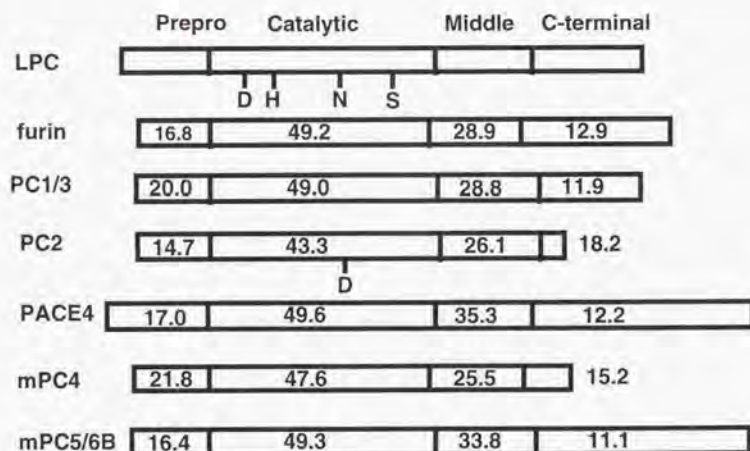
793

LPC. PRO LDVPHGKEEQIC.
MLPC. PRO PDL LQGKSGQIC.
RLPC. PRO PDL. . . KDGQI. C
XENOLPC. PRO TDLRHLTSEL. . .

5-2 Subtilisin-like serine protease

Subtilisin-like serine protease 遺伝子ファミリーは、近年報告されるようになった新しい遺伝子ファミリーである。すでに、PC1/3、PC2、PACE4、PC5/6、*furin* の6 遺伝子が報告されており、LPC は、7 番目のメンバーである。

他のメンバーと1 次配列を比較すると、LPC は他のメンバーを極めて異なっていることがわかる。以下に各ドメインごとの相同性を示す。



Numbers indicate sequence similarity with LPC

図37 LPC と他のメンバーとの比較

もっとも保存されている catalytic domain でも、保存されているアミノ酸残基が50%に満たない。

分子進化を phylogenetic tree で表すと、次のようになる⁶⁵。この結果も、LPC 遺

伝子が他の subtilisin-like protease とかなり異なっており、むしろ酵母の Kex2 遺伝子に近いことがわかる。これは、LPC 遺伝子が furin の活性中心付近の保存されている塩基配列を元にして degenerated PCR によっては、クローニングされなかったことと一致している。

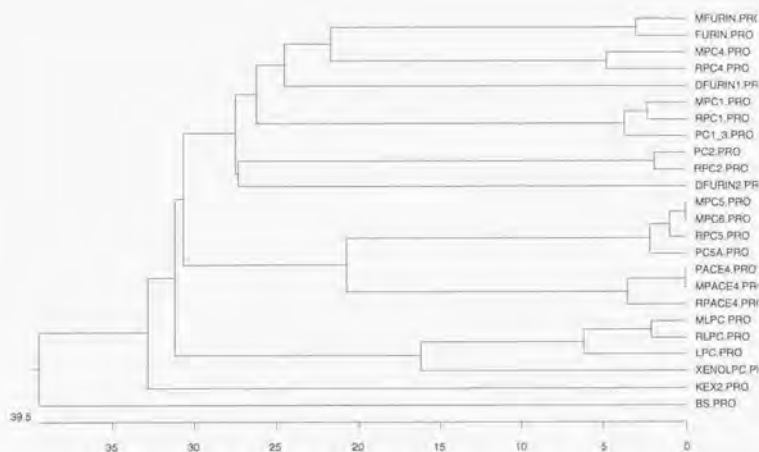


図38 subtilisin-like protease の Phylogenetic tree

この Phylogenetic tree から、LPC はかなり早い時期に他の subtilisin-like protease から分かれたと考えられる。それは、真核生物の diversity が大きくなった時期と考えられる。また *Xenopus* LPC 遺伝子の染色体構造が明らかになれば、より詳細な Phylogenetic tree が作成できると考えられる。

次に、subtilisin-like protease ファミリー間で保存されている活性中心領域の配列を他のメンバーと比較した結果を示す。

		186		227
MLPC	TGRGTVVVVDDGVET		NGNHHGTRCAGEIAA	
MFURIN	TGHGIVVSIILDDGLEK		NDNHHGTRCAGEVAA	
MPC1	TGKGVVITVLD DGLEW		NENKHHGTRCAGEIAM	
MPC4	TGRGVVISIILDDGIEK		DENKHHGTRCAGEVSA	
MPC5	TGKNIVVTIILDDGIER		NENKHHGTRCAGEVAA	
MPC6	TGKNIVVTIILDDGIER		NENKHHGTRCAGEVAA	
FURIN	TGHGIVVSIILDDGLEK		NDNHHGTRCAGEVAA	
LPC	TGRGTVVVVDDGVVH		NGNHHGTRCAGEIAA	
PC1_3	TGKGVVITVLD DGLEW		NENKHHGTRCAGEIAM	
PC2	TGKGVITIGIMDDGIDY		WFNSHHGTRCAGEVSA	
PC5A	TGKNIVVTIILDDGIER		NENKHHGTRCAGEVAA	
Bacillus	TGSNVKAVV DSGIDS		DNNSHHGTHVAGTVAA	
XENOPUS LPC	TGRGTVVVVDDGVQH		SDNHHGTRCAGEIAA	
Yeast KEX2	TGAGVVAAIYDDGLDY		SDDYHHGTRCAGEIAA	
		328		405
MLPC	GSIFVVASGNGG		TGTSAAAPLAAGMIALMLQ	
MFURIN	GSIFVWASGNGG		TGTSASAPLAAGIHALTLE	
MPC1	GSIFVWASGNGG		TGTSASAPLAAGIFALALE	
MPC4	GTLFIWASGNGG		TGTSASAPLAAGMIALALE	
MPC5	GSVFVWASGNGG		TGTSASAPMAAGIHALALE	
MPC6	GSVFVWASGNGG		TGTSASAPMAAGIHALALE	
FURIN	GSIFVWASGNGG		TGTSASAPLAAGIHALTLE	
LPC	GSIFVVASGNGG		TGTSAAAPLAAGMIALMLQ	
PC1_3	GSIFVWASGNGG		TGTSASAPLAAGIFALALE	
PC2	GSIFYWASGDGG		SGTSAAAPEAAGVFALALE	
PC5A	GSVFVWASGNGG		TGTSASAPMAAGIHALALE	
Bacillus	GVVVVAAAGNEG		NGTSMASPHVAGAAALILS	
XENOPUS LPC	GSIFVVASGNGG		TGTSAAAPPIAAGMIALMLQ	
Yeast KEX2	GAIYVVASGNGG		GGTSAAAPLAAGVYTLILLE	

図1 subtilisin-like protease の保存領域比較

図中の数字は、マウス LPC 配列中のアミノ酸残基を示す。また、活性中心であるアミノ酸残基は四角で囲んだ。この図から明らかなように、種を越えて活性中心およびその近傍のアミノ酸配列が保存されていることがわかる。このことは、これらの酵素の活性および基質が極めて似ていることを示唆している。in vitro 系の研究から、furin と PACE4 の両方が HIV の gp160 タンパク質のプロセッシングを行うことができることが報告されている。しかし、両者とも gp160 に対して、かなり低い酵素活性しか持たないことも報告されている⁶⁶。このことは、furin と PACE4、および少なくとももう一つ別の subtilisin-like protease が gp160 タンパク質を切断できることを示唆している。すなわち、in vitro 系では基質が重なりあうことを示している。

他の subtilisin-like protease の基質の検索もなされているが、in vitro 系では基質が重なりあうため、現在 native な基質を同定することは困難である。

subtilisin-like protease が切断する配列は、R-X-K/R-R、R-X-X-R、あるいは R/K-R である。

これまでに報告された基質は、以下通りである。

表26 subtilisin-like protease と基質

enzymes	furin	PC1/3, PC2
substrates	von Willebrand factor (vWF)	pro-opiomelanocortin (POMC)
	β -NGF	pro-insulin
	anthrax toxin	pro-glucagon
	HIV gp160	
	low density lipoprotein receptor-related protein (LRP)	

ただし, subtilisin-like protease 間で基質特異性は交差するため, HIV gp160 をプロセッシングする酵素として, furin 以外にも PACE4, そして LPC が候補となっている⁹⁾。

gp160 のプロセッシングを阻害することにより, HIV に感染した T 細胞の増殖を阻害できることが知られており, subtilisin-like protease の基質特異性の研究進展が期待されている。

LPC 遺伝子産物は, アミノ酸の 1 次配列の比較と解析から, furin と同様な立体構造をとることが予想されている。これは, furin 同様に 6 アミノ酸残基を認識している可能性を示唆するものである。現在, LPC 遺伝子産物の基質特異性に関しては, 次の表に示したことが知られている (personal communication)。

表27 Furin と LPC による基質の切断

		Endogenous	Furin	LPC
Pro-renin	DVRTKR	-	+	+/-
	DVFRTK	-	-	-
	RVRTKR	-	+	+
	RVFTQR	-	-	-
	RVRTQR	-	+	-
	RVFTKR	-	+	+
	DVRTKK	-	+/-	-
	DVFTKR	-	-	-
	DVRTQR	-	+/-	-
	DRRRKR	-	+	+/-
pro-vWF	GHRSKG	-	-	-
	GHRSKR	-	+	+
	GHR SAR	-	+	-
	GHASKR	-	+/-	-

上記の表から, LPC 遺伝子産物は, 切断部位 (示したアミノ酸配列の C 末端) から数えて, 1, 2, 4 番目のアミノ酸配列を認識していることがわかる。

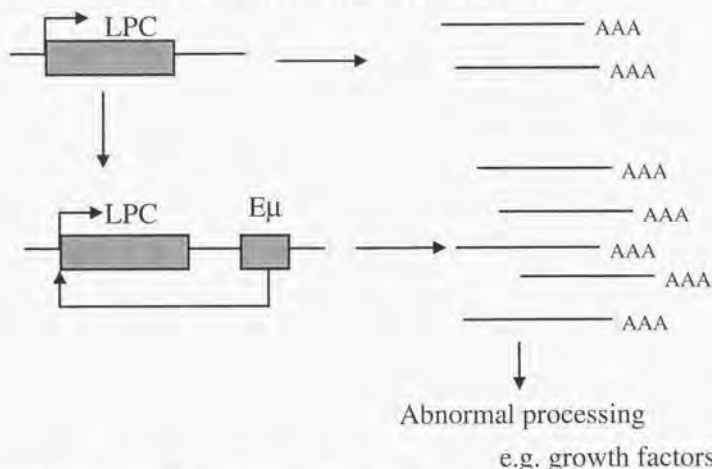
5-3 LPC と転座の関連

緒言で述べたように、ヒト LPC 遺伝子は白血病細胞において頻繁に転座が観察される 11q23 領域にマップされる。また、non-Hodgkin's リンパ腫細胞においてまれに見られる t(11;14)転座部位からクローニングされたため、転座およびリンパ腫との関連が示唆されている⁶⁾。

LPC 遺伝子が関与したリンパ腫の例は、今回の例以外に 2 例のみ知られているだけであるが、どちらも、LPC 遺伝子の 3' non-coding 領域に breakpoint が見つかっている。これは、この領域の転座がある種のリンパ腫の発症にとって何等かの役割をになっていることを示唆する。

白血病細胞で見られる転座と違って、融合タンパク質が発現することはないため、融合によってどのような変化がもたらされるかは不明である。しかし、転座によって、イムノグロブリンの E μ エンハンサーが LPC 遺伝子近傍に存在するようになることから、LPC 遺伝子の過剰発現が関与していることが考えられる (下図)。

LPC and lymphoma



あるいは、LPC 遺伝子の 3' non-coding 領域 (mRNA 内) が欠失することにより、LPC mRNA の安定性が変化し turn over が短くなることも考えられる。

LPC と同じファミリーに属する PC1/3 と疾病に関する報告によると、PC1 遺伝子の変異によって、肥満が引き起こされることが示唆されている。発症のメカニズムは、以下のように考えられている。POMC は、PC1/3 および PC2 によってプロセッシングされ、melanocortin が作られる。これが、melanocortin のリセプターである MC4-R を産生している細胞へと作用する。しかし、PC1/3 に変異がおこることにより、melanocortin が産生されず、正常な刺激が届かず肥満になると考えられる (下図参照)。このことは、MC4-R をノックアウトしたマウスが肥満になることと一致している。

Obesity and PCs

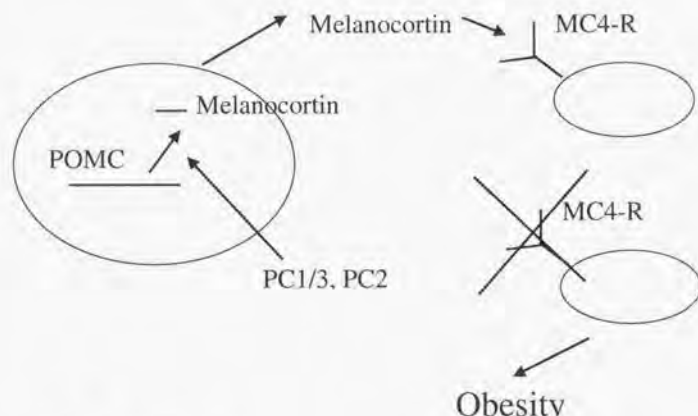


図 40 PC1/3 と肥満

5-4 今後の研究

LPC 遺伝子産物がリンパ腫の発症に関与している可能性を調べるために、現在、LPC 遺伝子を CD2 遺伝子のプロモーター下に結合したプラスミドを用いてトランスジェニックマウスを作製している。このベクターを用いることにより、リンパ細胞で LPC 遺伝子過剰発現を誘導することができる。また、LPC 遺伝子の機能を研究するために、knockout マウスを作成している。knockout マウスを作製することにより、mRNA の安定性の影響を間接的に解析できると共に、LPC の発達との関連が調べられるものと思われる。

6 謝辞

この研究を行うにあたりご指導いただいた Bryan D. Young 博士に感謝いたします。また、ヒト LPC 遺伝子の解析にあたり、Joanne Meerabux 博士および Nathalie Lecointe 博士に大変ご協力いただきました。

さらに、研究を進めるにあたってご助言いただいた英国 Imperial Cancer Research Fund, Medical Oncology の皆様に感謝します。

7 参考文献

- ¹ Young D.B. (1992) Cytogenetic and molecular analysis of chromosome 11q23 abnormalities in leukaemia. *Bailliere's Clinical Haematology* 5:881-185
- ² Oshimura M. et al. (1977) Chromosomes and causation of human cancer and leukaemia. XXVII Possible mechanisms for the genesis of common chromosome abnormalities, including isochromosomes and the Ph1. *Cancer Res.* 37:3501-3507
- ³ Van den Berghe H. et al. (1979) A new chromosome abnormality in acute lymphoblastic leukaemia (ALL). *Human Genet.* 46:173-181
- ⁴ Abe R and Sandberg A. A. (1984) Significance of abnormalities involving chromosomal segment 11q22-25 in acute leukaemia. *Cancer Genet. Cytogenet.* 13:121-127
- ⁵ Raimondi S. C. et al. (1989) Childhood acute lymphoblastic leukaemia with chromosomal breakpoint at 11q23. *Blood* 73:1627-1634
- ⁶ Milteman F. et al. (1993) Report of the committee on chromosome changes in neoplasia. *Genome Priority Reports* 1:700-726
- ⁷ Arthur D. C. et al. (1982) Translocation 4;11 in acute lymphoblastic leukaemia: clinical characteristics and prognostic significance. *Blood* 59:96-99

-
- ⁸ Kaneko Y. et al. (1986) Clinical and hematologic characteristics in acute leukaemia with 11q23 translocations. *Blood* 67:484-491
- ⁹ Hagemeyer A. et al. (1987) Characterization of the blast cells in acute leukemia with translocation (4;11): report of eight additional cases and of one case with a variant translocation. *Leukemia* 1:24-31
- ¹⁰ Heim B. et al. (1987) New structural chromosomal rearrangements in congenital leukemia. *Leukemia* 1:16-23
- ¹¹ Sait S. N. et al. (1987) A t(1;11) in acute nonlymphocytic leukemia FAB type M4. *Cancer Genet. Cytogenet.*24:181-183
- ¹² Feder M et al (1985) A2p:11q chromosome translocation in dysmyelopoietic preleukemia. *Cancer Genet. Cytogenet.*15:143-150
- ¹³ Derre J. et al. (1990) In situ hybridization ascertains the presence of a translocation t(6;11) in an acute monocytis leukemia. *Genes, Chromosomes and Cancer.*2:341-344
- ¹⁴ hagemeyer A. et al. (1982) Translocation (9;11)(p21;q23) in three cases of acute monoblastic leukemia. *Cancer Genet. Cytogenet.*5:95-105
- ¹⁵ Pui C. H. et al. (1987) An analysis of leukemic cell chromosomal features in infants. *Blood* 69:1289-1293
- ¹⁶ Hayashi Y et al, (1988) Immnogloblin heavy chain gene rearrangements and mixed

lineage characteristics in acute leukemias with the 11;19 translocation. *Cancer* 61:721-720

¹⁷ Bernard O. A. and Berger R. (1995) Molecular basis of 11q23 rearrangements in hematopoietic malignant proliferations. *Genes, Chromosome and Cancer* 13:75-85

¹⁸ Kaneko Y et al. (1988) Clinical characteristics of infant acute leukemia with or without 11q23 translocations. *Leukemia* 2:672-676

¹⁹ Katz F. et al. (1988) Cellulat and molecular studies on infant null acute lymphoblastic leukemia. *Blood* 71:1438-1447

²⁰ Stark B et al. (1989) Biologic and cytogenetic characteristics of leukemia in infants. *Cancer* 63:117-125

²¹ Chuu W. M. et al (1990) Infant leukemia: an analysis of mine Chinese patients. *American J. Hematology* 34:246-251

²² gibbons B. wt al (1990) Infant acute leukemia with t(11;19). *British J. Haematology* 74:264-269

²³ Rubnitz J. E. et al. (1994) Frequency and prognostic significance of HRX rearrangements in infant acute lymphoblastic leukemia: a Pediatric Oncology Group Study. *Blood* 84:570-573

²⁴ Sorensen P. H. et al.(1994) Molecular rearrangements of the MLL gene are present in

most cases of infant acute myeloid leukemia and are strongly correlated with monocytic or myelomonocytic phenotypes. *J. Clinical Investigation* 93:429-437

²⁵ Le Beau M. M. et al. (1986) Clinical and cytogenetic correlations in 63 patients with therapy-related myelodysplastic syndromes and acute nonlymphocytic leukemia: further evidence for characteristic abnormalities of chromosomes no5 and 7. *J. Clinical Oncology* 4:325-345

²⁶ Zulfian G. B. et al. (1993) Etoposide and secondary haematological malignancies: coincidence or causality? *Annals of Oncology* 4:559-566

²⁷ Ziemien-van der Poel S. et al. (1991) Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc. Natl. Aca. Sci. USA* 88:10753-10739

²⁸ Gu et al. (1992) The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila trithorax*, to the AF-4 gene. *Cell* 71:701-708

²⁹ Djabali M. et al. (1992) A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukemia. *Nature Genet.* 2:113-118

³⁰ Tkachuk D. C. et al. (1992) Involvement of a homolog of *drosophila-trithorax* by 11q23 chromosomal translocations in acute leukemias. *Cell* 71:691-700

³¹ Tschiersch B. et al. (1994) The protein encoded by the *Drosophila* position-effect

variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* 13:3822-3831

³² Grosschedl R. et al. (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends in Genetics* 10:94-100

³³ Prasad R. et al. (1993) Cloning of the ALL-1 fusion partner, the AF-6 gene, involved in acute myeloid leukemias with the t(6;11) chromosome translocation. *Cancer Res.* 53:5624-5628

³⁴ Bernard O. et al. (1994) A novel gene, AF-1F, fused to HRX in t(1;11)(p32;q23), is not related to AF-4, AF-9 or ENL. *Oncogene* 9:1039-1045

³⁵ Corral J. et al. (1993) Acute leukemias of different lineages have similar MLL gene fusions encoding related chimeric proteins resulting from chromosomal translocation. *Proc. Natl. Acad. Sci. USA.* 90:8538-8542

³⁶ Prasad R. et al. (1994) Leucine zipper dimerization motif encoded by the AF17 gene fused to ALL1(MLL) in acute leukemia. *Proc. Natl. Acad. Sci. USA.* 91:8107-8111

³⁷ Chaplin T. et al. (1995) A novel class of zinc finger/leucine zipper genes identified from the molecular cloning of the t(10;11) translocation in acute leukemia. *Blood* 85:1435-1441

³⁸ Thirman M. J. et al. (1994) Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia. *Proc. Natl. Acad. Sci. USA.* 91:12110-

12114

³⁹ Tse W. et al. (1995) A novel gene AF1q, fused to MLL in t(1;11)(q21;q23), is specifically expressed in leukemic and immature hematopoietic cells. *Blood* 85:650-656

⁴⁰ Schichman S. A. et al. (1994) ALL-1 partial duplication in acute leukemia. *Proc. Natl. Acad. Sci. USA.* 91:6236-6239

⁴¹ Gu Y. et al. (1994) Sequence analysis of the breakpoint cluster region in the ALL-1 gene involved in acute leukemia. *Cancer Res.* 54:2326-2330

⁴² Gu Y. et al. (1992) The 84;11(q21;q23) chromosome translocations in acute leukemias involve in the VDJ recombinase. *Proc. Natl. Acad. Sci. USA.* 89:10464-10468

⁴³ Steiner D. and Oyer P. E. (1967) The biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma. *Proc. Natl. Acad. Sci. USA.* 57:145

⁴⁴ Hutton J. C. (1990) subtilisin-like proteases involved in the activation of proproteins of the eukaryotic secretory pathway. *Curr. Op. Cell Biol.* 2:1131

⁴⁵ Hosaka M. et al. (1991) Arg-X1ys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. *J. Biol. Chem.* 266:12127

⁴⁶ Julius D. et al. (1984) Isolation of the putative structural gene for the lysine-arginine cleaving endopetidase required for processing of yeast prepro- α -factoe. *Cell* 37:1075

-
- ⁴⁷ Roebroek A. et al. (1986) Characterization of human fes/fps reveals a new transcription unit (fur) in the immediate upstream region of the proto-oncogene. *Mol Biol. Rep.* 11:117
- ⁴⁸ Schalken J. A. et al. (1987) fur gene expression as a discriminating marker in human lung cancer. *J. Clinical Investigation* 80:1545
- ⁴⁹ Dubois C.M et al. (1995). Processing of transforming growth-factor-beta-1 precursor by human furin convertase. *J Biol Chem.* 270: 10618-10624.
- ⁵⁰ Anderson-E.D et al.(1993). Inhibition of HIV-1 gp160-dependent membrane-fusion by a furin-directed alpha(1)-antitrypsin variant. *J Biol Chem.* 268: 24887-24891.
- ⁵¹ Inocencio N. M. et al. (1997) Endoprotease activities other than furin and PACE4 with a role in processing of HIV-1 gp160 glycoproteins in CHO-K1 cells. *J. Biol. Chem.* 272:1344-1348
- ⁵² Van de Van W. J. M. et al. (1990) Furin is a subtilisin-like proprotein processing enzyme in higher eukaryotes. *Mol. Biol. Rep.* 14:265
- ⁵³ Seidah N. G. et al. (1990) cDNA sequence of two distinct pituitary homologous to Kwx2 and furin gene products: Tissue-specific mRNAs encoding candidates for pro-hormone processing proteiase. *DNA Cell Biol.* 9:415-424
- ⁵⁴ Smeekens S.P. and Steiner D. F. (1990) Identification of a human insulinoma cDNA

encoding a novel mammalian protein structurally related to the yeast dibasic processing Kex2. *J. Biol. Chem.* 265:2997-3000

⁵⁵ Kiefer M. C. et al. (1991) Identification of a second human subtilisin-like protease gene in the *fes/fps* region of chromosome 15. *DNA Cell Biol.* 10:757-769

⁵⁶ Nakamura K. et al. (1992) Identification of a fourth member of the mammalian endoprotease family homologous to the yeast Kex2 protease. *J. Biol. Chem.* 267:5897-5900

⁵⁷ Seidah N. G. et al. (1992) Testicular expression of PC4 in the rat: molecular diversity of a germ cell-specific Kex3/subtilisin-like proprotein convertase. *Mol. Endocrinol.* 6:1559-1570

⁵⁸ Lusson J. et al. (1993). cDna structure of the mouse and rat subtilisin kexin-like pc5 - a candidate proprotein convertase expressed in endocrine and nonendocrine cells. *Proc Natl Acad Sci U S A.* 90: 6691-6695.

⁵⁹ Nakagawa T. et al. (1993). Identification of an isoform with an extremely large cys-rich region of pc6, a kex2-like processing endoprotease. *FEBS Lett.* 327: 165-171.

⁶⁰ Molloy S.S. et al. (1994). Intracellular trafficking and activation of the furin proprotein convertase - localization to the tgn and recycling from the cell-surface. *EMBO J.* 13: 18-33.

⁶¹ Meerabux J. M. et al. (1994) Molecular cloning of a novel 11q23 breakpoint

associated with non-Hodgkin's lymphoma. *Oncogene*. 9:893-898

⁶² Buckler A. J. et al. (1991) Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci. USA*. 88:4005-4009

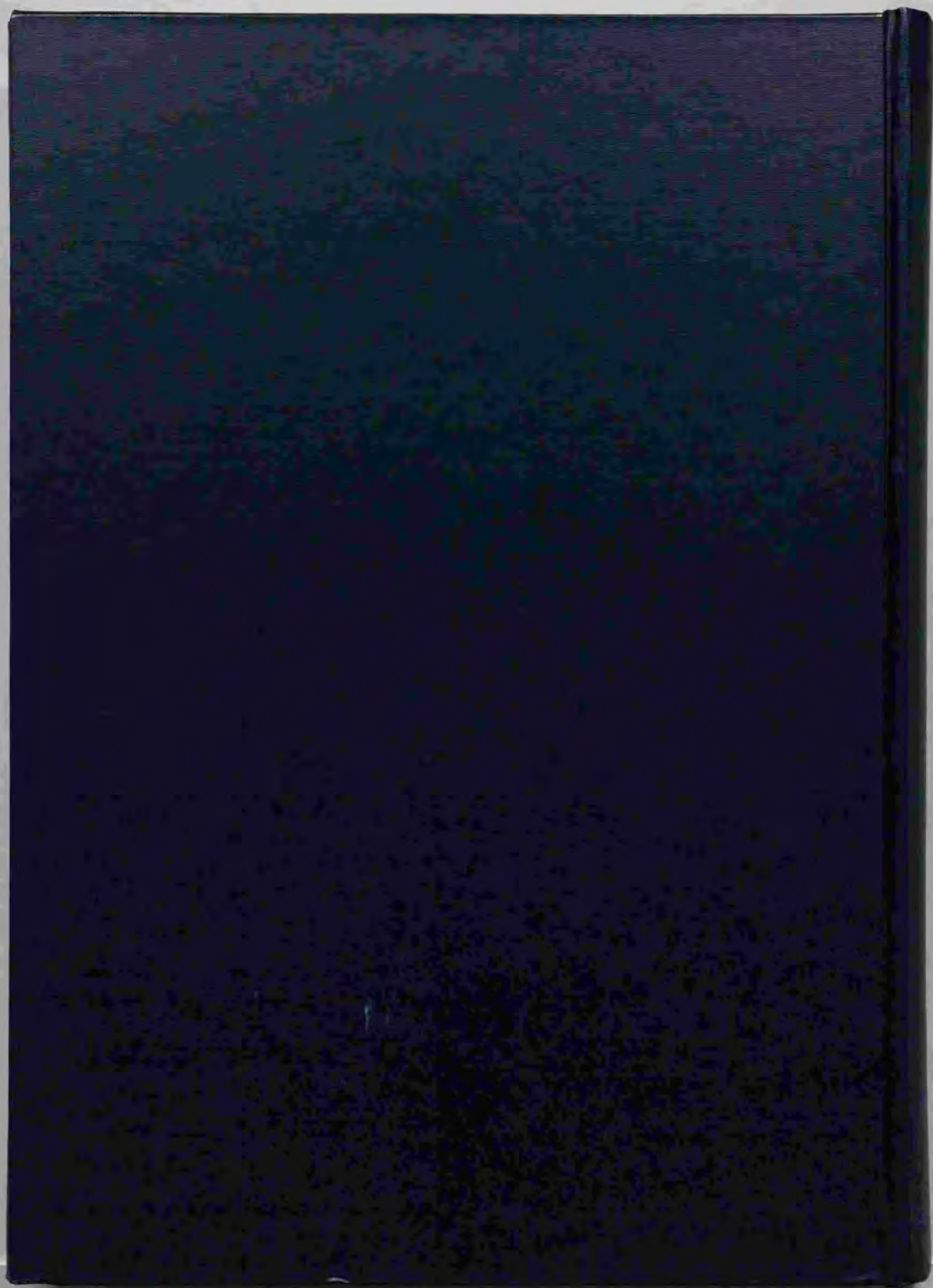
⁶³ Roebroek A J M et al. (1995). The *dfur2* gene of *drosophila-melanogaster* - genetic organization, expression during embryogenesis, and pro-protein processing activity of its translational product *dfurin2*. *DNA Cell Biol.* 14: 223-234.

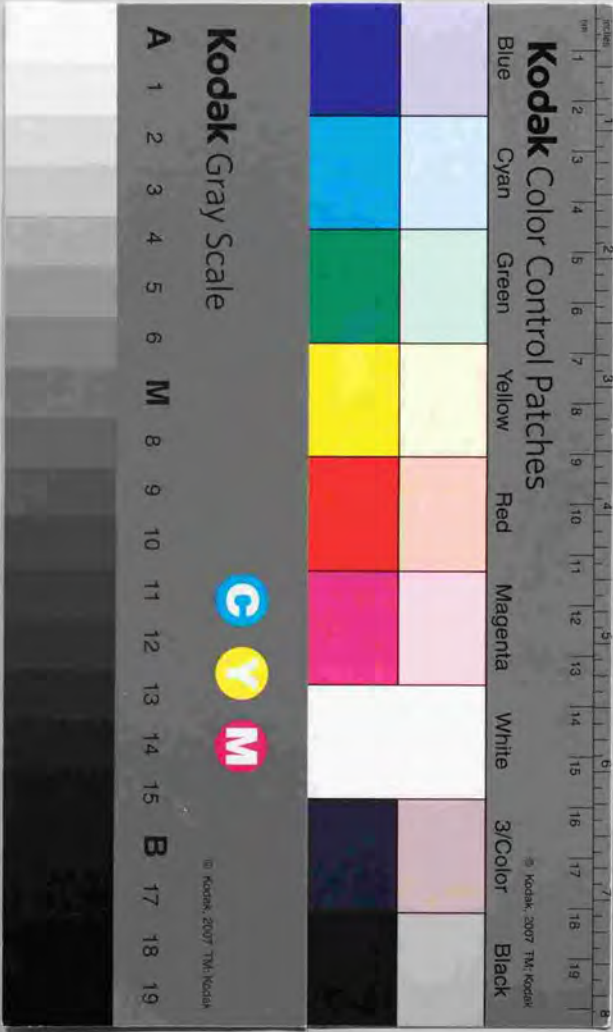
⁶⁴ Mbikay M et al. (1994). Structure of the gene for the testis-specific proprotein convertase-4 and of its alternate messenger-RNA isoforms. *Genomics*. 20: 231-237.

⁶⁵ Seidah N G et al. (1996). cDna structure, tissue distribution, and chromosomal localization of rat *pc7*, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases. *Proc Natl Acad Sci U S A*. 93: 3388-3393.

⁶⁶ Vollenweider F et al. (1996). Comparative cellular processing of the human-immunodeficiency-virus (HIV-1) envelope glycoprotein gp160 by the mammalian subtilisin/kexin-like convertases. *Biochem J*. 314: 521-532.

⁶⁷ Meerabux J et al.(1996). A new member of the proprotein convertase gene family (*lpc*) is located at a chromosome-translocation breakpoint in lymphomas. *Cancer Res.* 56: 448-451.





Kodak Color Control Patches

Blue Cyan Green Yellow Red Magenta White 3/Color Black

Kodak Gray Scale

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19



© Kodak, 2007 TM Kodak