TUMSAT-OACIS Repository - Tokyo University of Marine Science and Technology (東京海洋大学)

4倍性魚種コイのc-myc遺伝子2タイプの進化

メタデータ	言語: jpn 出版者: 公開日: 2008-03-31 キーワード (Ja): キーワード (En): 作成者: 二見, 邦彦 メールアドレス: 所属:
URL	https://oacis.repo.nii.ac.jp/records/754

4倍性魚類コイの c-*myc* 遺伝子 2タイプの進化

平成14年度

(2002)

東京水産大学大学院

水産学研究科

資源育成学専攻

二見邦彦

目次

博士論文内容の要旨……………………………………………………………………………… 1

緒論……………………………………………………………………………………………… 3

第１章　研究史……………………………………………………………………………………7

第１節　c-*myc*遺伝子について……………………………………………………… 8

第２節　ゲノムの倍数性と生物進化…………………………………………………… 15

第２章　コイ c-*myc*遺伝子の転写開始点の決定…………………………………… 25

第３章　コイ *Max* の cDNA クローニングと、c-*myc* との発現の比較…………43

第４章　コイ c-*myc*遺伝子２タイプの機能分化…………………………………… 68

第５章　Bacterial two-hybrid system を用いた新規 c-Myc 結合タンパク質の網羅的スクリーニング……………………………………………………… 101

総合考察……………………………………………………………………………………… 113

謝辞…………………………………………………………………………………………… 117

資料　総説　４倍性魚類コイの c-*myc*遺伝子２タイプの進化……………… 118

〔課程博士〕

博士論文内容の要旨（No.1）

（2,000 字程度）

報告番号	課博第 号	氏　名	二見 邦彦

（要　旨）

論文題目：　「4 倍性魚類コイの c-*myc* 遺伝子 2 タイプの進化」

生物の進化において遺伝子の重複は極めて重要な過程であり、特にゲノムの倍数性は、5 億年前のカンブリア期における脊椎動物の爆発的進化（カンブリア爆発）の主役であったと考えられている。魚類における倍数性の解明は、脊椎動物の進化と遺伝子重複との関係の謎を解く鍵である。ヒトを始め、多くの動物で倍数性は知られているが、高等脊椎動物での倍数性化はかなり昔（数億年前）に起きたもので、重複した遺伝子間の相同性が低くなり、いわゆる 4 倍性の 2 倍性化が進んでいるため、重複した遺伝子の進化を解明することは困難である。しかし、4 倍性魚類であるコイは比較的最近（数千万～1 億年前）ゲノムの倍数性化が起きたため、二つの遺伝子間の相同性はまだ高い。したがって、倍数性化後の遺伝子の進化を研究する上で、ゲノムの倍数化したコイは格好のモデルになるといえる。そのコイではすでに 2 タイプの c-*myc* 遺伝子がクローニングされており（c-*myc*1 と c-*myc*2）、しかも、その両方が発現していることが明らかにされている。

c-*myc* は、細胞の増殖、分化の抑制、アポトーシスの誘導などに関わる重要な遺伝子であり、その異常は細胞の癌化を引き起こす。したがって、倍数性化後も少なくとも一つの遺伝子は必ず保存的である必要があるので、もう一方の遺伝子の進化を保存的な遺伝子の機能と比較して解析を進めることができる。

これまで、遺伝子の構造から倍数性進化について議論した報告はいくつかあるものの、遺伝子発現や機能の側面から研究した報告は少ない。そのため、倍数性進化の仮説に実証を与えるまでには至っていない。そこで本研究では、c-*myc* 遺伝子の発現調節機構および機能に関する研究を行い、遺伝子機能の分化の側面からゲノムの倍数性化による進化の仮説に実証を与えることを目的とした。

<u>コイ c-*myc* 遺伝子 2 タイプの転写開始点の決定</u>

コイ c-*myc* 遺伝子の転写開始点をオリゴキャップ法により解析し、魚類 c-*myc* のエキソン 1 の存在を初めて明らかにした。c-*myc*1 と c-*myc*2 のエキソン 1 の間の相同性は他のエキソンに比べて低く、転写開始点の位置や数も異なっていた。このことから、c-*myc*1 と c-*myc*2 は異なる発現制御機構を持つ可能性が示唆された。

<u>コイ *Max* の cDNA クローニングおよび c-*myc* 遺伝子 2 タイプとの発現の比較</u>

RT-PCR により各組織での c-*myc* の発現量を定量したところ、発現様式に違いが認められた。またゼブラフィッシュでは、c-*myc* が Myc 蛋白質と二量体をつくる Max の遺伝子の発

博士論文内容の要旨（No.2）

現と共同歩調をとることが分かっているため、*Max* の cDNA をクローニングし、ノーザンブロットを行ったところ、*Max* は c-*myc2* と同様の発現様式を示したが c-*myc1* とは異なった。系統解析から c-*myc1* は c-*myc2* より進化速度が速いことが分かっており、このことから c-*myc1* は c-*myc* としての機能とは別の新しい機能を持つように進化している可能性が推測された。

<u>コイ c-*myc* 遺伝子 2 タイプの機能分化</u>

コイの c-*myc* 遺伝子 2 タイプの間にどのような機能の分化が起きているのかを明らかにするために、培養細胞株を用いて c-Myc 蛋白質の生化学的機能の違いについて解析をおこなった。まず、ニシキゴイ由来の培養細胞株である KF と KG の細胞周期をそれぞれ G_0 に同調させ、血清刺激後の c-*myc* の mRNA 量を RT-PCR で定量したところ、KF では c-*myc2* の発現が誘導されたのに対し、KG では逆に c-*myc1* の発現が誘導された。そこで、遺伝子発現制御に影響を与えると考えられる転写開始段階での機能の違いを明らかにするために、c-*myc* の 5'上流領域の欠失変異体を作製し、ルシフェラーゼアッセイを行ったところ、c-*myc1* と c-*myc2* は異なる転写制御機構を持つことが分かった。

次に、c-Myc と Max の蛋白質間相互作用を調べるために、GST-pull down assay および two-hybrid system による解析を行ったところ、進化速度が速い c-Myc1 は、Max との結合が c-Myc2 よりも弱く、*in vitro* ではほとんど結合しなかった。

また、ゲルシフトアッセイの結果では、両 c-Myc はその標的配列である E-box に対して親和性に違いは見られなかったものの、レポーター遺伝子を用いての cotransfection assay では、保存的な c-Myc2 は c-Myc1 に比べて高い転写活性をもつ傾向が見られた。*in vivo* では、c-Myc の標的遺伝子のうち、*TERT* の発現は c-Myc2 によってのみ制御されていた。

<u>Bacterial two-hybrid system によるコイ c-Myc 結合蛋白質の網羅的スクリーニング</u>

それぞれの c-Myc と特異的に相互作用する蛋白質を網羅的にスクリーニングするために、bacterial two-hybrid system をおこなったところ、KF cDNA ライブラリーから c-Myc1 と c-Myc2 に対してそれぞれ異なる 11 個の候補蛋白質が得られ、相互作用する蛋白質においても 2 つの c-Myc の間に違いが認められた。

これらの結果から、c-*myc1* と c-*myc2* の機能はオーバーラップしてはいるものの、この 2 つは明らかに区別できるものであった。確証を与えることはできなかったが、ゲノムの倍数性化後、進化速度の速い c-*myc1* は発現する組織や細胞を変え、種々の結合蛋白質との組み合わせで標的遺伝子を変化させている可能性がある。このことは、魚類においても、哺乳類にみられるような *myc* ファミリーが形成されつつあることを想定させる。

緒言

1970 年代に 大野 乾 は、生物の進化において遺伝子の重複は極めて重要であり、多細胞生物の多様化、組織の複雑化に密接に関わっているという仮説を提唱した（Ohno, 1970）。遺伝子重複には部分的重複と倍数性とがある。一つの遺伝子が重複すると、二つになったうちの一つは新しい機能を持つように進化する自由が与えられるが、倍数性化は、染色体上のすべての遺伝子座が重複するため、構造遺伝子と調節遺伝子のバランスが保たれる。そして調節遺伝子の分岐により、構造遺伝子の発現の組織特異性が向上する。しかし、部分的重複では必ずしも常に調節遺伝子が含まれるとは限らず、そのため倍数性化は、部分的重複とは異なる重要性を進化の過程で持っている。ゲノムの倍数性は、植物の進化においては重要な働きを果たしてきたことが知られている他、脊椎動物においても約 5 億年前のカンブリア期におけるその爆発的進化（“カンブリア爆発”、“進化のビッグバン”などとも呼ばれる）の主役であったと考えられている。

魚類における倍数性の解明は、脊椎動物の進化と遺伝子重複との関係の謎を解く鍵であると考えられている。ヒトを始め、多くの動物で倍数性は知られているが、致死作用があったり、たとえ生存しても健全性を欠くものが多い。そ

れに対し、魚類は現在でも倍数性化の能力を保持しており、異なる属間個体も比較的容易に交配でき、染色体の可塑性を保持している（小島、1983）。さらに、高等脊椎動物での倍数性化はかなり昔（数億年前）に起きたもので、重複した遺伝子間の相同性が低くなり、いわゆる 4 倍性の 2 倍性化が進んでいるため、重複した遺伝子の進化を解明することは困難である。しかし、魚類におけるゲノムの倍数性化は比較的最近（数千万～1 億年前）起きたため、二つの遺伝子間の相同性はまだ高い（Ohno, 1970）。したがって、倍数性化後の遺伝子の進化を研究する上で、ゲノムの倍数化した魚類は格好のモデルになるといえる。魚類においては、サケ科やコイ科で多くの倍数性化の例が報告されており、実際コイから 2 タイプの c-*myc* 遺伝子がクローニングされており、しかも、その両方が発現していることが明らかにされている（Zhang *et al.*, 1995）。

c-*myc* は細胞の増殖、分化の抑制、アポトーシスの誘導などに関わる重要な遺伝子で、その異常が細胞の癌化を引き起こすことから癌遺伝子としても知られている（for review, see Oster *et al.*, 2002）。したがって、倍数化後も少なくとも一つの遺伝子は必ず保存的である必要があるため、もう一方の遺伝子の進化を保存的な遺伝子の機能と比較して解析を進めることができる。また、ヒトゲノムには、5 種類のメンバーからなる *myc* ファミリー（c-*myc*、L-*myc*、N-*myc*、s-*myc*、B-*myc*）が報告されているが、これらの起源、類縁関係はいまだに不明

である。倍数性化により重複したコイの *c-myc* 遺伝子の研究は、ヒトにおける *myc* ファミリーの形成の解明にも役立つかもしれない。

これまで、遺伝子の構造から倍数性進化について議論した報告は数多くあるものの、遺伝子発現や機能の側面から研究した報告は少ない。そのため、倍数性進化の仮説に実証を与えるまでには至っていない。そこで本研究では、*c-myc* 遺伝子の発現調節機構および機能に関する研究を行い、遺伝子機能の分化の側面からゲノムの倍数性化による進化の仮説に実証を与えることを目的とした。

文献

Ohno, S.: Evolution by Gene Duplication. Springer Verlag Press, Heidelberg, New York, 1970

小島吉雄：魚類細胞遺伝学．水交社，東京，1983

Oster, S. K., Ho, C. S., Soucie, E. L. and Penn, L. Z.: The *myc* oncogene: MarvelouslY Complex. *Adv. Cancer Res.*, **84**, 81-154 (2002)

Zhang, H., Okamoto, N. and Ikeda, Y.: Two c-*myc* genes from a tetraploid fish, the common carp (*Cyprinus carpio*). *Gene*, **153**, 231-236 (1995)

第1章　研究史

第１節　c-*myc* 遺伝子について

歴史と概要

核局在性癌遺伝子 *myc* は最初、ニワトリ骨髄細胞腫ウイルス 29（myelo-cytomatosis virus29,MC29）というレトロウイルスで同定され、v-*myc* と呼ばれた。MC29 ウイルスは、骨髄細胞腫、上皮性の悪性腫瘍、肉腫、およびリンパ腫を引き起こす（Cole, 1986)。また、他のニワトリレトロウイルス、例えば MH2、CMⅡ、OK10 や、ネコ白血病ウイルスにも *myc* が挿入されている（Burck *et al.*, 1988)。

c-*myc*（cellular *myc*）は v-*myc* の細胞性ホモログであり、ヒトでは第 8 染色体長腕 24 領域にある（Dalla-Favera *et al.*, 1982)。今までに、ヒトの染色体では、c-*myc* 遺伝子と共通の構造的、機能的特徴を持つ遺伝子がいくつか存在することが報告されている（Ryan *et al.*, 1996)。それらは c-*myc*、N-*myc*、L-*myc* の 3 種類の重要な遺伝子の他に、s-*myc*、B-*myc* といった計 5 種類のメンバーによる *myc* ファミリーを構成している。各遺伝子の構造は似ているものの、塩基配列およびアミノ酸配列の相同性は 60％以下である。これらの起源、類縁関係は不明である。

myc ファミリーのうち、c-*myc*、N-*myc*、および L-*myc* の 3 種類の主要な遺伝子は、ヒトの癌で最も頻繁に活性化の見られる癌遺伝子である。例えば、c-*myc* はバーキットリンパ腫、急性リンパ性白血病、胃癌、肺癌、および大腸癌など多くの癌でその遺伝子の増幅が見つかっており、N-*myc* は神経芽細胞腫、肺小細胞癌に、L-*myc* は肺癌でのみ活性化が見つかっている（田矢、1991)。特に c-*myc* はヒト癌

の約3分の1で変異を起こしており（Cole, 1986)、現在最もよく研究されている癌遺伝子のひとつである。

このような経緯から、アンチセンス *myc* RNA を用いた癌の治療法の開発が行われており、また c-Myc にはアポトーシス誘導能が備わっていることから（後述)、これを利用した癌細胞の殺傷方法が考え出され、脳腫瘍の治療などへの適用が試みられている（Asai *et al.*, 1994)。

c-*myc* 遺伝子の構造と転写制御機構

哺乳類における c-*myc* 遺伝子の構造解析によれば、3つのエクソンと2つのイントロンの存在が知られている（Battey *et al.*, 1993)。転写開始点は4ヶ所（P0、P1、P2 および P3）存在する。主要な転写開始点は2カ所（P1 と P2）であるが、いずれの臓器、細胞種においても P2 からの転写産物の方が多い。一部の癌細胞で、染色体転座によりエクソン1が失われた場合には、イントロン1内の隠れたプロモーター（P3）から転写が開始される例がある。P0 からの転写については、その意義は不明である。Poly(A)付加位置も 2 ヶ所あり（pA1、pA2)、大部分の mRNA は pA2 で終わる（Hayashi *et al.*, 1987)。

c-*myc* 遺伝子の発現調節機構については、哺乳類ではよく研究されており、非コーディング領域であるエクソン 1 がその制御をしていることが明らかにされている。それは、ヒトのバーキットリンパ腫やマウスの形質細胞腫において、c-*myc* 遺伝子のエクソン 1 やその近辺の非コーディング領域に欠損や点変異がしばしばみられることから裏付けられる（Burck *et al.*, 1988)。この場合、c-*myc* が免疫グロ

ブリンやT細胞レセプターの遺伝子と相互転座を起こし、免疫グロブリン、T細胞レセプター遺伝子の強いエンハンサーの作用で c-*myc* が過剰発現する。

培養細胞では、増殖サイクルにある状態では c-*myc* の発現は E2F や Fos/Jun などの転写制御因子により正および負に調節されてことが指摘されている（林、1992)。しかし、哺乳類の c-*myc* 遺伝子では、第1イントロンにもいくつかのタンパク質結合部位が存在しており（Zajac-Kaye and Levans, 1990)、この部位で変異あるいは欠失が起きると c-*myc* の発現をコントロールできなくなる（Yu *et al.*, 1993）ことなどから、統一的な解釈は困難である。

通常 c-*myc* の発現量は低く（Zimmerman *et al.*, 1986)、このような遺伝子の定常的発現量決定には mRNA の半減期の長短も極めて重要である。c-*myc* の mRNA の代謝回転は極めて速く（半減期約 10〜30 分)、またその速度は、細胞の生理的条件によっても変動する（Dani *et al.*, 1984; Dani *et al.*, 1988)。c-*myc* のこのような不安定性は、poly(A)付加位置近傍に存在する A+U-rich 領域が関係している（Pei and Calame, 1988; Brewer and Ross, 1988)。また、核 run-on アッセイを用いた実験で、transcription pausing によって全長を持つ mRNA の発現低下がもたらされるという結果も報告されている（Bentley and Groudine, 1986; Nepveu and Marcu, 1986)。

c-Myc タンパク質の構造と機能

SDS-PAGE では、ヒトの c-Myc タンパク質は 64 kDa と 67 kDa の少なくとも二つが観察される（Nau *et al.*, 1985)。後者は第1エクソンの 3'末端付近に存在す

る CUG（Leu）から翻訳されるが、その生物学的な意義は不明である。尚、Met 以外から始まる哺乳動物細胞遺伝子の翻訳としては c-Myc の例がはじめてである。

Myc タンパク質は C 末端領域に塩基性領域－ヘリックス・ループ・ヘリックス－ロイシンジッパードメイン（bHLH-LZ）といった典型的な転写因子としての構造をもち、パートナーの Max (Myc Associate Protein X)とロイシンジッパーにおいて結合し、ヘテロダイマーを形成する（Blackwell *et al.*, 1990; Blackwood and Eizenman, 1991)。それによって DNA 上の CACGTG 配列（E-box）の認識とそれへの結合が達成され、転写活性化因子として機能する。Max のホモダイマーも Myc/Max と同様の塩基配列に結合するが、転写活性化能はなく、Myc/Max と拮抗し転写抑制に働くものと考えられている。Myc/Max ヘテロダイマーは Myc による形質転換とアポトーシスの両方に極めて重要であるが、最近、N 末端領域に結合する転写コアクチベーターTRRAP など、Max 以外の Myc 結合タンパク質もいくつか報告されてきている（Oster *et al.*, 2002)。

細胞の癌化に重要な c-Myc タンパク質の標的遺伝子は複数同定されており（Grandri and Eisenman, 1997; Cole and McMahon, 1999, Oster *et al.*, 2002)、直接細胞の癌化に関わると思われるのは、ODC、eIF4E および CDC25A であろうと言われている（有賀ら、1997)。また最近、テロメラーゼ遺伝子のプロモーター領域にも Myc の結合部位が見つかり、Myc がテロメラーゼを活性化することもわかってきた（Wang *et al.*, 1998; Wu *et al.*, 1999)。しかし、Myc は活性化する遺伝子の組み合わせによって、まったく異なる生物効果を細胞に与える可能性があり、そのため、Myc によって直接制御される遺伝子を明らかにすることは、非常に難しいと考えられている（Peters and Vousden, 1997)。

生体内での正常機能

c-myc は PDGF などによる増殖刺激で一過性に過剰発現することなどから細胞周期における G_0/G_1 移行の制御に深く関与し(Kelly *et al.*, 1983; Greenberg *et al.*, 1984)、特に、細胞の増殖、分化の抑制、アポトーシス、トランスフォーメーション、ゲノムの不安定性、血管新生などに重要な役割を果たす（Oster *et al.*, 2002)。実際、*c-myc* のノックアウトマウスは胎児期に死亡し、また *c-myc* を過剰発現するトランスジェニックマウスは発現に使うエンハンサーの種類に応じて乳癌やリンパ腫などを発生する（Henriksson and Lüscher, 1996)。一方、栄養因子除去という特定の条件下で、*c-myc* の強制発現はアポトーシスを誘導させることも明らかとなり（Askew *et al.*, 1991; Evan *et al.*, 1992)、その標的遺伝子としては癌抑制遺伝子である *p53* などが知られている（Hermeking and Eick, 1994; Wagner *et al.*, 1994)。これらのことから *c-myc* は、細胞のホメオスタシス維持に重要な役割を果たしていることがわかってきた。そのため、哺乳類(Hayashi *et al.*, 1987; Bernard *et al.*, 1983; Watt *et al.*, 1983)、鳥類（Watson *et al.*, 1983)、両生類（King *et al.*, 1986)、魚類(Van Beneden *et al.*, 1986; Schreiber-Agus *et al.*, 1993; Zhang *et al.*, 1994; Zhang *et al.*, 1995)、原索動物（二見ら、未発表)、および棘皮動物（Walker *et al.*, 1992）の間で同遺伝子はよく保存されており、ショウジョウバエ（Villares and Cabrera, 1987）やカキ（Marsh and Chen, 1995）でも類似の配列がクローニングされている。

c-*myc* に関する研究は数多く報告されているが、シグナル伝達経路は未解決な部分が多く、c-*myc* の発現や機能の制御は十分に理解されているとは言い難い（Lusche *et al.*, 2001)。その理由は、c-*myc* を発現していない細胞株が存在しない（c-*myc*(-)にすると細胞は死んでしまう）ため、外来からの導入 c-*myc* 遺伝子を用いての機能研究が難しいためといわれている（有賀ら、1997)。今後、c-Myc タンパク質と複合体形成する蛋白質がつぎつぎに cDNA クローニングされることで、これらの問題は解決されると考えられる。

魚類におけるこれまでの c-*myc* 遺伝子の研究

魚類では c-*myc* 遺伝子がニジマス（Van Beneden *et al.*, 1986)、ゼブラフィッシュ（Schreiber-Agus., 1993)、コイ（Zhang *et al.*, 1995)、キンギョ（Zhang *et al.*, 1994）でクローニングされている。ニジマスでは、イントロン 2 とエクソン 2、3 はすでに明らかにされており、エクソン 3 がよく保存されているほか、エクソン 2 の中の Box A と B の領域も特によく保存されている。エクソン 1 領域の存在は魚類ではまだ明らかにされていない。

4 倍性魚類であるコイでは 2 タイプの c-*myc*（*CAM*1 と *CAM*2）がクローニングされ、mRNA も約 2.1 kb と約 1.5 kb の 2 種類が確認されている（巻末の資料 図 1、2 参照)。それらは異なる連鎖群に属しており（Sun ら，私信)、さらにそれぞれ進化速度が異なり、*CAM*1 は *CAM*2 よりも進化速度が 1.6 倍速いことが分かっている（張、1994)。それらのアミノ酸配列はヒトの c-Myc と高い相同性をもち、それぞれ 55.3%と 56.7%である。倍数化によって生じた 2 タイプの c-*myc* 遺伝子

である *CAM1* と *CAM2* の間の相同性は 94.2%であり、その値からコイにおける 4 倍性化は約 5,800 万年前に起こったと推定されている（Zhang *et al.*, 1995）。魚類における c-*myc* と腫瘍との関係に関する研究はほとんどない（岡本ら，1997）。

第2節　ゲノムの倍数性と生物進化

Ohno（大野）の仮説

生物の進化において、遺伝子重複は極めて重要である。特にゲノムの倍数性は、植物の進化では重要な働きを果たしてきたことが知られているが、近年、脊椎動物の進化においても重要な過程と考えられるようになった。1960 年代末に Ohno（大野）らは、哺乳類と鳥類の細胞当たりの DNA 量が魚類や他の脊索動物より多く、また数多くの遺伝子座が重複していることから、「脊椎動物の進化過程に、少なくとも一回のゲノム倍数性化が起きた」と提唱した。実際に、約 5 億年前のカンブリア期に脊椎動物の共通祖先に一回の 4 倍性化が起きたと推測している（Ohno, 1970; Ohno, 1967; Ohno, 1968)。さらにゲノム解析技術の発展につれ、哺乳類、特にヒトやマウスにおいては、原始的な脊索動物から哺乳類までの間に 3 回のゲノム倍数化が起きたことが示唆された（Lundin *et al.*, 1993)。

一般的に、倍数性は約 5 億年前におきた爆発的進化（カンブリア爆発）の主役であったと考えられている。一つの遺伝子が重複すると、二つになったうちの一つは新しい機能を持つように進化する自由が与えられるが、倍数性化はすべての遺伝子が重複するため、構造遺伝子と調節遺伝子のバランスが保たれる。調節遺伝子の分岐により、構造遺伝子発現組織の特異性が向上する（Ohno, 1970)。

ゲノムの倍数性と魚類

遺伝子重複には部分的重複とゲノムの倍数性とがある（Ohno, 1970）。部分的重複では、すでに成り立っている遺伝的システムに多少の修飾が加えられるに留まるか、もしくはすでにある遺伝的システムに阻害的な働きをするかもしれない。それに対し、ゲノムの倍数性は、システム全体が重複するため、すでに存在している遺伝的システムを阻害することなく、新たな体制をもった遺伝的システムが構築される可能性がある。そのため倍数性化は、部分的重複とは異なる重要性を進化の過程で持っている。しかし、ヒトを始め、多くの動物で倍数性は知られているが、致死作用があったり、たとえ生存しても健全性を欠くものが多い。さらに、高等脊椎動物での倍数性化はかなり昔（数億年前）に起きたもので、重複した遺伝子はすでに違った機能を獲得したり、偽遺伝子化したりしたため、塩基配列の相同性が低くなり、いわゆる 4 倍性の 2 倍性化が進んでいる。そのため、重複した遺伝子の進化を解明するのは困難である（Ohno, 1993）。それに対し、性決定機構が確立されていない魚類や両生類は今でも倍数性化の能力を保持しており、異なる属の間でも容易に交雑でき、染色体の可塑性を保持している（Becak *et al.*, 1966; 小島、1983; Turner, 1984）。

最初の脊椎動物 4 倍体種として発見されたのはブラジル産のカエル（*Odontophrynus americanus*）であった（Becak *et al.*, 1966）。しかし、この種の 4 倍体化はつい最近起こったもので、4 つの相同染色体があり、減数分裂時にこの 4 つの染色体が互いに対合して、二つの 2 価染色体でなく一つの 4 価染色体を形成する（同質 4 倍体）。4 倍体化した瞬時では、遺伝子重複現象は起こらない。なぜなら、既存の遺伝子座それぞれが、旧来の 2 対立遺伝子の代わりに 4 対立遺伝子を得

た状態となっただけだからである。長い年月を経て初めて、1遺伝子座から2遺伝子座が独立し、それぞれがまた2対立遺伝子をもつ状態に戻れるわけである。したがって、ブラジル産カエルに続いて見つけたかったのは、それよりもっとずっと以前に4倍体化した種であった。

コイは数千万〜1億年前に倍数化が起きたとされており、同じコイ科の他の魚の染色体数が約50であるのに対し、104の染色体数を持つことから、4倍性魚類であると考えられている。コイの104の染色体は減数分裂中26の4価染色体を作らず、52個の2価染色体を作る。すなわちこれは重複した遺伝子座の一つ一つが独立して2個の遺伝子座になっていることを示している。実際、コイにおいてはすでに c-*myc* のほか、補体C3（Nakao, *et al.*, 2000)、ミオシン重鎖（Kikuchi *et al.*, 1999）など多くの遺伝子が複数クローニングされており、その両方が発現していることが明らかにされている。したがって、コイは倍数化後の重複遺伝子の進化を研究する上で格好のモデルとなり得るわけで、コイにおける倍数性の解明は、脊椎動物の進化と遺伝子重複との関係の謎を解く鍵であると考えられている（大野、1996)。現在、魚類はヒトを含めた脊椎動物の研究に不可欠な実験系となっており（武田ら，2000)、コイの c-*myc* 遺伝子の研究は脊椎動物における *myc* ファミリーの形成の解明など発癌研究にも役立つかもしれない。

文　献

有賀寛芳・有賀早苗：*myc* がん遺伝子ファミリー．“がん遺伝子研究の展望II（野田亮　編)”．「*現代化学*」増刊，**33**，18-24 (1997)

Asai, A., Miyagi, Y., Hashimoto, H., Lee, S. H., Mishima, K., Sugiyama, A., Tanaka, H., Mochizuki, T., Yasuda, T., Kuchino, Y.: Modulation of tumor immunogenicity of rat glioma cells by s-Myc expression: eradication of rat gliomas *in vivo*. *Cell Growth Differe.*, **5**, 1153-1158 (1994)

Askew, D. S., Ashmun, R. A., Simmons, B. C. and Cleveland, J. L.: Constitutive c-*myc* expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis. *Oncogene*, **6**, 1915-1922 (1991)

Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenior, G. and Leder, P.: The human c-*myc* oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. *Cell*, **34**, 779-787 (1983)

Becak, M. L., Becak, W. and Rabello, M. N.: Cytological evidence of constant tetraploidy in the bisexual South American frog Odontophrynus americanus. *Chromosoma*, **19**, 188-193 (1966)

Bentley, D. and Groudine, M.: A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. *Nature*, **321**, 702-706 (1986)

Bernard, O., Suzanne, C., Steven, G., Elizabeth, W. and Jerry, M. A.: Sequence of the murine and human cellular *myc* oncogenes and two modes of *myc* transcription resulting from chromosome translocation in B lymphoid tumors. *EMBO J.*, **2**, 2375-2383 (1983)

Blackwell, T. K., Kretzner, L., Blackwood, E.M., Eisenman R. N. and Weintraub H.: Sequence-specific DNA binding by the c-Myc protein. *Science*, **250**, 1149-1151 (1990)

Blackwood E. M. and Eizenman R. N.: Max: A Helix-Loop-Helix Zipper Protein That Forms a Sequence-Specific DNA-Binding Complex with Myc. *Science*, **251**, 1211-1217 (1991)

Brewer, G and Ross, J.: Poly(A) shortening and degradation of the 3' A+U-rich sequences of human c-*myc* mRNA in a cell-free system. *Mol. Cell. Biol.*, **8**, 1697-1708 (1988)

Burck, K. B., Liu, E. T., Larrick, J. W.: Oncogenes; An Introduction to the Concept of Cancer Genes. Springer-Verlag Press, New York, 1988

Cole, M. D.: The *myc* oncogene: its role in transformation and differentiation. *Annu. Rev. Gen.* **20**, 361-384 (1986)

Cole, M. D. and McMahon S. B.: The Myc oncoprotein: a critical evaluation of transactivation and target gene regulation. *Oncogene*, **18**, 2916-2924 (1999)

Dalla-Favera, R., Gelmann, E. P., Martinotti, S., Franchini, G., Papas, T. S., Gallo R. C. and Wong-Staal F.: Cloning and characterization of different human sequences related to the onc gene (v-*myc*) of avian myelocytomatosis virus (MC29). *Proc. Natl. Acad. Sci. USA*, **79**, 6497-6501 (1982)

Dani, C., Blanchard, J-M., Piechaczyk, M., El Sabrouty, S., Marty, S and Jeanteur, P.: Extreme instability of *myc* mRNA in normal and transformed human cells. *Proc. Natl. Acad. Sci. USA*, **81**, 7045-7050 (1984)

Dani, C., Mechti, N., Piechaczyk, M., Lebleu, B., Jeanteur, P. and Blanchard, J-M.: Increased rate of degradation of c-*myc* mRNA in interferon-treated Daudi cells. *Proc. Natl. Acad. Sci. USA*, **82**, 4896-4899 (1985)

Evan, G. I., Wyllie, A. H., Gilbert, C. S., Littlewood, T. D., Land, H., Brooks, M., Waters, C. M., Penn L. Z. and Hancock D. C.: Induction of apoptosis in fibroblasts by c-*myc* protein. *Cell*, **69**, 119-128 (1992)

Grandri, C. and Eisenman, R. N.: Myc target genes. *Trends. Biochem. Sci.*, **22**, 177-181 (1997)

Greenberg, M. E. and Ziff, E. B.: Stimulation of 3T3 cells induces transcription of the c-*fos* proto-oncogene. *Nature*, **311**, 433-438 (1984)

Hayashi, K., Makino, R., Kawamura, H., Arisawa, A. and Yoneda, K.: Characterization of rat c-*myc* and adjacent regions. *Nucleic Acid research*, **15**, 6419-6436 (1987)

林健志：c-*myc* 遺伝子．*蛋白質核酸酵素*，**37**，860-861 (1992)

Henriksson, M. and Lüscher, B.: Proteins of the Myc network: essential regulators of cell growth and differentiation. *Adv. Cancer Res.*, **68**, 109-182 (1996)

Hermeking, H., Eick, D.: Mediation of c-Myc-induced apoptosis by p53. *Science*, **265**, 2091-2093 (1994)

Kelly, K., Cochran, B. H., Stiles C. D. and Leder P.: Cell-specific regulation of the c-*myc* gene by lymphocyte mitogens and platelet-derived growth factor. *Cell*, **35**, 603-610 (1983)

Kikuchi, K., Muramatsu, M., Hirayama, Y. and Watabe, S.: Characterization of the carp myosin heavy chain multigene family. *Gene*, **228**, 189-196 (1999)

King, M. W., Roberts, J. M. and Eisenman, R. N.: Expression of the c-*myc* proto-oncogene during development of *Xenopus laevis*. *Mol. Cell. Biol.*, **12**, 4499-4508 (1986)

Lundin, L. G.: Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics*, **16**, 1-19 (1993)

Luscher, B.: Function and regulation of the transcription factors of the Myc/Max/Mad network. *Gene*, **277**, 1-14 (2001)

Marsh, A. G. and Chen, T. T.: A divergent cDNA homologue of the c-*myc* proto-oncogene in the eastern oyster Crassostrea virginica: implications for Myc evolution. *Mol. Mar. Biol. Biotechnol.*, **4**, 185-192 (1995)

Nakao, M., Mutsuro, J., Obo, R., Fujiki, K., Nonaka, M. and Yano, T.: Molecular cloning and protein analysis of divergent forms of the complement component C3 from a bony fish, the common carp (*Cyprinus carpio*), presence of variants lacking the catalytic histidine. *Eur. J. Immunol.*, **30**, 858-866 (2000)

Nau, M. M., Brooks, B. J., Battey, J., Sausville, E., Gazdar, A. F., Kirsch, I. R., McBride, O. W., Bertness, V., Hollis G. F. and Minna J. D.: L-*myc*, a new *myc*-related gene amplified and expressed in human small cell lung cancer. *Nature*, **318**, 69-73 (1985)

Nepveu, A and Marcu, K.: Intragenic pausing and anti-sense transcription within the murine c-*myc* locus. *EMBO J.*, **5**, 2859-2865 (1986)

Ohno, S.: Sex Chromosomes and Sex-Linked Genes. Springer Verlag Press, Heidelberg, New York, 1967

Ohno, S., Wolf, U. and Atkin, N. B.: Evolution from fish to mammals by gene duplication. *Hereditas*, **59**, 169-187 (1968)

Ohno, S.: Evolution by Gene Duplication. Springer Verlag Press, Heidelberg, New York, 1970

Ohno, S.: Patterns in genome evolution. *Curr. Opin. Genet. Dev.*, **3**, 911-914 (1993)

大野 乾：続 大いなる仮説—5.4 億年前の進化のビッグバン．羊土社，東京，1996

小島吉雄：魚類細胞遺伝学．水交社，東京，1983

岡本信明・正仁親王：がん遺伝子とがん抑制遺伝子．“魚類の DNA（青木宙・隆島史夫・平野哲也 編）．pp. 414-423，恒星社厚生閣，東京，1997

Oster, S. K., Ho, C. S., Soucie, E. L. and Penn, L. Z.: The *myc* oncogene: MarvelouslY Complex. *Adv. Cancer Res.*, **84**, 81-154 (2002)

Pei, R. and Calame, K.: Differential stability of c-*myc* mRNAS in a cell-free system. *Mol. Cell. Biol.*, **8**, 2860-2868 (1988)

Peters, G. and Vousden, H.: Oncogenes and Tumor Suppressors. Oxford University Press, London, 1997

Ryan, K. M. and Birnie, G. D.: *Biochem. J.*, *Myc* oncogenes: the enigmatic family. **314**, 713-721 (1996)

Schreiber-Agus, N., Horner, J., Torres, R., Fung-Chow Chiu and Depinho, R. A.: Zebra Fish *myc* and *max* Genes: Differential Expression and Oncogenic Activity throughout Vertebrate Evolution. *Mol. Cell. Biol.*, **13**, 2765-2775 (1993)

武田洋幸・岡本仁・成瀬清・堀寛：小型魚類研究の新展開.「*蛋白質核酸酵素*」12月号増刊，**45**，p.316 (2000)

田矢洋一：*myc* と情報伝達. *実験医学*, ９, 169-172 (1991)

Turner, B. J.: Evolutionary Genetics of fishes. Plenum Press, New York and London, 1984

Van Beneden, R. J., Watson, D. K., Chen, T. T., Lautenberger, J.A. and Papas, T. S.: Cellular *myc* (c-*myc*) in fish (rainbow trout): its relationship to other vertebrate *myc* genes and to the transforming genes of the MC29 family of viruses. *Proc. Natl. Acad. Sci. USA*, **83**, 3698-3702, (1986)

Villares, R. and Cabrera, C. V.: The achaete-scute gene complex of D. melanogaster: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. *Cell*, **50**, 415-424 (1987)

Wagner, A. J., Kokontis, J. M. and Hay, N.: Myc-mediated apoptosis requires wild-type p53 in a manner independent of cell cycle arrest and the ability of p53 to induce p21waf1/cip1. *Genes Dev.*, **8**, 2817-2830 (1994)

Walker, C. W., Boom, J. D. G. and Marsh, A. G.: First non-vertebrate member of the *myc* gene family is seasonally expressed in an invertebrate testis. *Oncogene,* **7**, 2007-2012 (1992)

Wang, J., Xie, L. Y., Allan, S., Beach, D. and Hannon, G. J.: Myc activates telomerase. *Genes and Development*, **12**, 1769-1774 (1998)

Watson, D. K., Reddy, E. P., Duesberg, P. H. and Papas, T. S.: Nucleotide sequence analysis of the chicken c-*myc* gene reveals homologous and unique coding regions by comparison with the transforming gene of avian myelocytomatosis virus MC29, delta gag-*myc*. *Proc. Natl. Acad. Sci. USA*, **80**, 2146-2150 (1983)

Watt, R,. Nishikura, K., Sorrentino, J., ar-Rushdi, A., Croce, C. M. and Rovera, G.: The structure and nucleotide sequence of the 5' end of the human c-*myc* oncogene. *Proc. Natl. Acad. Sci. USA*, **80**, 6307-6311 (1983)

Wu, K-J., Grandri, C., Amacker, M., Simon-Vermot, N., Polack, A., Lingner, J. and Dalla-Favera, R.: Direct activation of TERT transcription by c-MYC. *Nature Genetics*, **21**, 220-224 (1999)

Yu, B. W., Ichinose, I., Bonham, M. A. and Zajac-Kaye, M.: Somatic mutations in c-*myc* intron I cluster in discrete domains that define protein binding sequences. *J. Biol. Chem.*, **268**, 19586-19592 (1993)

Zajac-Kaye, M. and Levans, D.: Phosphorylation-dependent binding of a 138-kDa *myc* intron factor to a regulatory element in the first intron of the c-*myc* gene. *J. Biol. Chem.*, **265**, 4547-4551 (1990)

張寰：4倍性コイ科魚類、コイ、フナの c-*myc* 遺伝子の単離とそれを用いた系統進化．東京水産大学博士論文、1994

Zhang, H., Zhang, H., Okamoto, N. and Ikeda, Y.: Cloning and Sequencing of a c-*myc* Gene from Goldfish *Carassius auratus*. *Fisheries Science*, **60**, 707-711 (1994)

Zhang, H., Okamoto, N. and Ikeda, Y.: Two c-*myc* genes from a tetraploid fish, the common carp (*Cyprinus carpio*). *Gene*, **153**, 231-236 (1995)

Zimmerman, K. A., Yancopoulos, G. D., Collum, R. G., Smith, R. K., Kohl, N. E., Denis, K. A., Nau, M. M., Witte, O. N., Toran-Allerand, D., Gee, C. E., Minna, J.D., Alt, F. W.: Differential expression of *myc* family genes during murine development. *Nature*, **319**, 780-783 (1986)

第2章　コイ c-*myc*遺伝子の転写開始点の決定

Futami K, Komiya T, Zhang H and Okamoto N: Determination of heterogeneous transcription start points of two c-*myc* genes from the common carp (*Cyprinus carpio*).

Gene, **245**, 43-47 (2000)

Determination of heterogeneous transcription start points of two c-*myc* genes from the common carp (*Cyprinus carpio*)

Keywords: oligo-capping; the first exon; tetraploid; "TATA-less" promoter

Kunihiko Futami [a], Takeru Komiya [a], Huan Zhang [b], Nobuaki Okamoto [a,*]

[a] *Department of Aquatic Biosciences, Tokyo University of Fisheries, Konan 4, Minato-ku, Tokyo 108-8477, Japan*

[b] *IRAGO Institute, 377 Ehima-Shinden, Atsumi-cho, Aichi 441-3605, Japan*

*Corresponding author. Tel: +81 3 5463 0547, Fax: +81 3 5463 0552; e-mail: nokamoto@tokyo-u-fish.ac.jp

Abbreviations: bp, base pair(s); *CAM*1, carp c-*myc*1 encoding c-MYC1; *CAM*2, carp c-*myc*2 encoding c-MYC2; cDNA, DNA complementary to RNA; c-*myc*, gene(s) encoding c-MYC; c-MYC, cellular MYC; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; PCR, polymerase chain reaction; *tsp*, transcription start point(s).

ABSTRACT

We determined the heterogeneous transcription start points (*tsp*) of two c-*myc* genes from the common carp (*Cyprinus carpio*), tetraploid teleost, by the oligo-capping method and showed the existence of the first exon. This is the first report on the existence of the first exons of fish c-*myc* gene. Transcription of the two carp c-*myc* genes started from at least four sites in *CAM*1, locating from –752 to –381 bp upstream of the translation start site, and from twelve sites in *CAM*2, locating from -586 to -413 bp upstream respectively. The first intron of *CAM*1 and *CAM*2 were deduced to be 335 bp and 356 bp, respectively. They shared 86.9% nt identity, lower than those of the second exons (94.1%), and third exons (92.3%), which suggest that the first exon are evolving faster. No nt identities were found between the c-*myc* first exons of carp and other vertebrates. The putative promoter regions in *CAM*1 and *CAM*2 contained no obvious TATA or CCAAT boxes in the expected positions.

1. Introduction

Polyploidy is a potentially important process in the evolution of vertebrates (Ohno, 1970; Lundin, 1993). Studies on gene duplication in tetraploid teleosts are important for investigating the evolutionary processes following the tetraploid event (Ohno, 1993).

The proto-oncogene c-*myc* is thought to be one of the most important genes in controlling cell proliferation (Roy et al., 1993). It has precise expression (both specifically and quantitatively), is crucial for cell division and differentiation and is highly conserved in vertebrates. In mammals, c-*myc* genes consist of three exons and two introns (Bernard et al., 1983). The first exon is a noncoding exon. It plays a regulatory role in the transcription of the c-*myc* gene (Saito et al., 1983). The second and third exon together encode the c-MYC protein. It has been reported that the first exon evolved more quickly than the second and third exons (Bernard et al., 1983; Hayashi et al., 1987). Furthermore, the human c-*myc* gene is transcribed by two promoters (Bernard et al., 1983; Battey et al., 1983). The noncoding exon and the promoter structures have not been reported in the lower vertebrates.

Two c-*myc* genes in a tetraploid fish have been isolated from the common carp, *Cyprinus carpio* (Zhang et al., 1995). However, the first exon in the carp c-*myc* was not detected because of the incompleteness of the cloned carp c-*myc* cDNA. In addition, no signal was observed when carp genomic DNA was analyzed by Southern hybridization using human exon1 as a probe (Zhang et al., 1993). This result suggests that either exon1 is not present in carp c-*myc*, or it does exist, but its nt sequences are too different to be detected by human exon1 probe. Therefore, in this study, we determined the transcription start points (*tsp*) of two c-*myc* genes from the common carp by the

oligo-capping method (Maruyama and Sugano, 1994) and demonstrated the existence of the first exon.

2. Materials and methods

2.1. Isolation of RNA

Due to the fact that c-*myc* was detected in the liver of rainbow trout (Van Beneden et al., 1986) and carp, by preliminary experiment, hepatopancreas was selected for RNA extraction. Extraction was performed using Trizol reagent (Gibco BRL), according to the manufacturer's protocol. Poly(A)$^+$ mRNA was purified by Oligotex™-dT30 <Super> (TaKaRa).

2.2. Oligo-capping

Oligo-capping was performed as described by Maruyama and Sugano (1994) with some modifications. 5 μg of poly(A)$^+$ mRNA was briefly treated with bacterial alkaline phosphatase (BAP; TaKaRa). After two extraction with phenol:chloroform and ethanol precipitation, the cap structure of this poly(A)$^+$ mRNA was removed by tobacco acid pyrophosphatase (TAP HG; Nippon Gene, Toyama, Japan). After phenol:chloroform extraction and ethanol precipitation, the decapped mRNA was recapped with a chimeric RNA/DNA oligo linker (5'-GAG AGA GAC AGG CCT TGT TGG CCG AGA GG-3', 3'-ribose) using T4 RNA ligase (TaKaRa).

2.3. RT-PCR

The first strand cDNA of this oligo-capped mRNA was synthesized with RNaseH-free reverse-transcriptase (SUPERSCRIPT II, Gibco BRL) using an

oligo$(dT)_{12-18}$ primer. The PCR reaction was performed in a volume of 100 μl using AmpliTaq Gold and 10 $\times$ PCR Gold buffer (Perkin-Elmer Cetus, USA) with 0.2 μM of PCR primers. Zhang et al. (1995) cloned two c-*myc* genes *CAM*1 and *CAM*2, and determined the nt sequence of their 5' upstream regions. Using these nt sequences, c-*myc* specific primers were designed in the 5' upstream regions. These primers were oligo linker-specific primer (LSP-1, 5'-GAG AGA GAC AGG CCT TGT TGG CCG A-3'), and two c-*myc*-specific primers (Myc-1A, 5'-GTC CTT GCT GAT GGT GAT AGA AAT C-3' and Myc-2A, 5'-GTG ACA GAG GCA GGG TGA ATA-3'). PCR Amplification was started with a 12 min hold at 95°C, followed by 35 cycles of 30 s at 94°C, 1 min at 55°C, and 1 min at 72°C with a post-extension of 3 min at 72°C. Since a high level of nonspecific amplification was generated, the expected fragments were confirmed by Southern blot hybridization (Sambrook et al., 1989).

2.4. Nested PCR

The 5'-end of c-*myc* mRNA was amplified using LSP-2 primer (5'-ACA GGC CTT GTT GGC CGA GAG-3') with Myc-1B primer (5'-ATA CGC CAA ACT CGA ACT CAC CGG-3') or with Myc-2B primer (5'-ATA AAT TCT GTA GCT CCC GCG-3'). The reaction conditions are described in section *2.3.*

2.5. Subcloning and sequence analysis

The amplified fragments were separated by an agarose gel electrophoresis , and cloned into a home-made T-tailed pBluescript II SK(-) vector. Sequences were

analyzed by dye terminator cycle sequencing using the ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Cetus, USA).

3. Results and discussion

3.1. Determination of tsp *of two c-*myc *genes*

Identification of the 5' end of an mRNA is essential for determing the promoter region of a gene, especially when it is a "TATA-less" promoter region, where the *tsp* are not easily predictable. The oligo-capping method was used here to determine the *tsp* of two c-*myc* genes from the common carp.

PCR products amplified using primer LSP-2 with Myc-1B were cloned and the sequences of 21 positive clones were determined. Using the oligo-capping method, *tsp* of *CAM*1 were located at -466, -458 and -381 bp upstream of the putative translation initiation codon (The "A" of ATG is numbered as +1.), while *tsp* of *CAM*2 were located at -586, -581, -579, -569, -543, -525, -517, -494, -492, -484, -419 and -413 bp upstream (Fig. 1). Six of the 21 clones were found starting from -581 of *CAM*2. The sequence at -580 to -574 of *CAM*2 was homologous to the initiator element (PyPyANA/TPyPy). The major *tsp* in *CAM*2 may be at around -581.

Generally, it is difficult to amplify long fragments in PCR. So PCR was also performed using primers LSP-2 with Myc-2B which was designed in the upper stream regions. Of 12 clones that were sequenced, all were found to start at -752bp upstream of ATG of *CAM*1 (data not shown).

The sizes of the first introns interrupting the noncoding exon1 and exon2 were deduced to be 335 bp in *CAM*1 and 356 bp in *CAM*2, beginning with GT and ending with AG, respectively (Fig. 2). It is common that TATA box is around 25 bp upstream of *tsp*, and CCAAT box is around 80 bp upstream of *tsp*. However, characterization of the 5'-flanking regions indicated that the putative promoter regions in *CAM*1 and

CAM2 contained no obvious TATA or CCAAT box like sequences in the expected positions, which is different from mammalian c-*myc* genes. Certainly, the *CAM1* promoter contained two TATA sequences (nt –1233, -1200), while there are three in *CAM2* (nt -1389, -1254 and -904). However, as these sequences are situated far from any *tsp*, location of the promoters remains to be determined.

3.2. Comparison of the tsp *of* CAM*1 and* CAM2

The oligo-capping method indicated that transcription of the two c-*myc* genes of carp started from at least four sites in *CAM1* and from twelve sites in *CAM2*. In human and mouse, the c-*myc* gene has two *tsp* (Bernard et al., 1983; Battey et al., 1983). Our results showed more variable clones and raised the possibility of the presence of variable *tsp* in carp c-*myc* genes, since the oligo-capping method specifically labels the capped end of mRNAs. Indeed, using the oligo-capping method, variable *tsp* are obtained from human EF-1 α and TGF-β type II receptor genes (Maruyama and Sugano, 1994; Yu et al., 1996; Suzuki et al., 1997). The maximum distance between these *tsp* was 371 bp in *CAM1* and 173 bp in *CAM2* (Fig. 2). The difference in the *tsp* locations may be related to the modulation of expression.

Using "GENETYX-MAC" computer algorithm developed by Software Development Co., the nt identities of *CAM1* and *CAM2* were 78.5% in intron1, 86.9% in the first exon (nt –560 to –342 of *CAM1* and –586 to -364 of *CAM2*), 94.1% in the second exon, and 92.4% in the third exon. These results suggested that the first exon evolved faster than the second and third exon, which corresponds to the reports of Bernard et al. (1983) and Hayashi et al. (1987). Using the BLAST (Altschul et al., 1990) program, there are no nt identities between the c-*myc* exon1s of carp and other

vertebrates. In mammalian, several protein (MIF-1, MIF-2 and MIF-3) binding sites located at c-*myc* intron1 were identified and are controlling c-*myc* expression (Zajac-Kaye and Levens., 1990; Yu, B. W., 1993). But in carp c-*myc*, these protein binding sites were not observed. Therefore, the c-*myc* genes of carp may have a transcription regulation system that is different from that of other vertebrates. Indeed, high expression of the lower vertebrates c-*myc* in differentiated tissues contrasts sharply with the low levels observed in mammalian adult tissues (Schreiber-Agus et al., 1993; Schreiber-Agus et al., 1993). These differences may correlate with lower vertebrate-specific functions, such as tissue regeneration and/or immortalization of cell lines.

The tetraploid event has been recognized as an important process in the evolution of vertebrates (Ohno, 1970; Lundin, 1993; Ohno, 1993). The present study helps us to understand the transcription function and evolution of c-*myc* genes in tetraploid fishes as well as in other vertebrates, besides knowing the differences between the two c-*myc* genes. It is suggested that subsequent to the tetraploidization event, one of the 2 duplicated genes may have evolved faster to obtain a new function or become silent (Ohno, 1970). The differences in exon1 and the promoter structure between the two c-*myc* genes of carp suggested that *CAM*1 and *CAM*2 were evolving to acquire different functions after the tetraploid event. Further studies are needed to determine whether the exon1 of the carp c-*myc* genes plays a different role from that of other vertebrates.

3.3. Conclusions

(1) We determined the heterogeneous *tsp* of two carp c-*myc* genes by the oligo-capping method and indicated the existence of exon1. There are no nt identities between the c-*myc* first exons of carp and other vertebrates.

(2) The first exons of the carp c-*myc* genes are evolving faster than the second and third exons, which corresponds to the reports of Bernard et al. (1983) and Hayashi et al. (1987).

(3) Characterization of the 5'-flanking regions indicated that the putative promoter regions in *CAM*1 and *CAM*2 contain no obvious TATA or CCAAT boxes in the expected positions.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E., Lipman, D. J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G., Leder, P., 1983. The human c-*myc* oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell 34 779-787.

Bernard, O., Cory, S., Gerondakis, S., Webb, E., Adams, J. M., 1983. Sequence of the murine and human cellular *myc* oncogenes and two modes of *myc* transcription resulting from chromosome translocation in B lymphoid tumors. EMBL J. 2, 2375-2383.

Hayashi, K., Makino, R., Kawamura, H., Arisawa, A., Yoneda, K., 1987. Characterization of rat c-*myc* and adjacent regions. Nucleic Acids Res. 15, 6419-6436.

Lundin, L. G., 1993. Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16, 1-19.

Maruyama, K., Sugano, S., 1994. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. Gene 138, 171-174.

Ohno, S. (1970). Evolution by gene duplication. Springer Verlag Press, Heidelberg, New York.

Ohno, S., 1993. Patterns in genome evolution. Current Opinion in Genetics and Development 3, 911-914.

Roy, A. L., Carruthers, C., Gutjahr, T., Roeder, R. G., 1993. Direct role for Myc in transcription initiation mediated by interactions with TFII - I . Nature 365, 359-361.

Saito, H., Hayday A. C., Wiman, K., Hayward, W. S., Tonegawa, S., 1983. Activation

of the c-*myc* gene by transcription: A model for translation control. Proc. Natl. Acad. Sci. USA 80, 7476-7480.

Sambrook, J., Fritsch, E. F., Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schreiber-Agus, N., Horner, J., Torres, R., Chiu, F-C., DePinho, R. A., 1993. Zebra Fish *myc* and *max* Genes: Differential Expression and Oncogenic Activity throughout Vertebrate Evolution. Mol. Cell. Biol. 13, 2765-2775.

Schreiber-Agus, N., Torres, R., Horner, J., Lau, A., Jamrich, M., DePinho, R. A., 1993. Comparative Analysis of the Expression and Oncogenic Activities of *Xenopus* c-, N-, and L-*myc* Homologs. Mol. Cell. Biol. 13, 2456-2468.

Suzuki Y., Yoshitomo-Nagasawa, K., Maruyama, K., Suyama, A., Sugano, S. 1997. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200, 149-156.

Van Beneden, R. J., Watson, D. K., Chen, T. T. Lautenberger, J. A., Papas, T. S., 1986. Cellular *myc* (c-*myc*) in fish (rainbow trout): its relationship to other vertebrate *myc* genes and to the transforming genes of the MC29 family of viruses. Proc. Natl. Acad. Sci. USA 83, 3698-3702.

Yu, B. W., Ichinose, I., Bonham M. A., Zajac-Kaye, M., 1993. Somatic Mutation in c-*myc* Intron Ⅰ Cluster in Discrete Dominant That Define Protein Binding Sequences. J. Biol. Chem. 268, 19586-19592.

Yu, Y. S., Suzuki Y., Yoshitomo, K., Muramatsu, M., Yamaguchi, N., Sugano, S., 1996. The promoter structure of TGF-β type Ⅱ receptor revealed by "oligo-capping" method and deletion analysis. Biochem. Biophys. Res. Commun. 225, 302-306.

Zajac-Kaye, M., Levens, D., Phosphorylation-dependent Binding of a 138-kDa *myc*

Intron Factor to a Regulatory Element in the First Intron of the c-*myc* Gene. J. Biol. Chem. 265, 4547-4551.

Zhang, H., Okamoto, N., Ikeda, Y., 1995. Two c-*myc* genes from a tetraploid fish, the common carp (*Cyprinus carpio*). Gene 153, 231-236.

Zhang, H., Okamoto, N., Yamamoto, N., Ikeda, Y., 1993. Molecular cloning of carp cellular *myc* (c-*myc*) cDNA. Gyobyo kenkyu (Fish pathology) 28, 111-117.

Figure legends

Fig. 1. The 5' end sequences of the oligo-capped cDNA of two carp c-*myc* genes. The sequences corresponding to the carp c-*myc* genes were aligned along with the genomic sequences shown above. Dots (.) in *CAM*2 indicate the same residues as in *CAM*1. Clones 1 to 18 and 19 to 21 correspond to *CAM*2 and *CAM*1 respectively. Gaps (-) shown between the sequence derived from the linker oligo and the sequence corresponding to the carp c-*myc* genes do not exist in the real sequence. Six clones start from –581 bp upstream of the translation start site of *CAM*2

Fig. 2. The nt sequences of the 5' upstream regions of *CAM*1 and *CAM*2. Sequences in intron1 are indicated by lowercase letters, and other sequences are indicated by capital letters. Dots (.) in *CAM*2 indicate the same residues as in *CAM*1. Gaps (-) are introduced to optimize identity. The nt residues are numbered at the right. The putative translation start codon ATG is indicated by bold letters. Primers used in oligo-capping are indicated by horizontal arrows. *Tsp* are indicated by vertical arrows. The initiator (Inr)-like sequence is underlined. The nt sequences of *CAM*1 and *CAM*2 have DDBJ accession numbers of D37887 and D37888, respectively.

```
                 -569                                                                                          -474
CAM1       ------ACTGACACCCAGGCTATCAGTCTGTCCAGCACCTGAGGGGTGCTCGAGTGCGTACGTCGATTACGGTTTCCATCACTTTCTATTTGGCCAGGACTC
CAM2       ------T...T...A...A..ı.........T.....T.....A...A.ATAG.A...AG...T......AT..............C.....A.........
                 -596           -581                                                                           -501
clone 1                        AGACTATCAGTCTGTTCA------
      2                           ATCAGTCTGTTCAGCATC------
      3                           ATCAGTCTGTTCAGCATC------
      4                           ATCAGTCTGTTCAGCATC------
      5                           ATCAGTCTGTTCAGCATC------
      6                           ATCAGTCTGTTCAGCATC------
      7                           ATCAGTCTGTTCAGCATC------
      8                             CAGTCTGTTCAGCATCTG------
      9                             CAGTCTGTTCAGCATCTG------
     10                                   AGCATCTGAGAGGTACAT------
     11                                                    GCGTTGATTACATTTTCC------
     12                                                               ATCACTTTCCATTTGACC------
     13                                                                    CCATTTGACCAGGA---
          -473                                                                                         -375
CAM1       TGATTCCACTTATGCTGCAAAAGACCGGAGTAAACTATTGGATCTTACGTTTTGTTCCTTATTTTTTGCTCTCTTTTCATTGCATTGCGTCTCGAACGA---
CAM2       ..............T.....G...............G..........T.....C....G........T..........C.....C.........GA..---
          -500                                                                                         -402
clone14        CACTTATGTTGCAAGAGA------
     15          CTTATGTTGCAAGAGACC------
     16               GCAAGAGACCGGAGTAAA------
     17                                                                       GCATCGCGTCTCGAGA---
     18                                                                          CGTCTCGAGAGA---
     19          CTTATGCTGCAAAAGACC------
     20               GCAAAAGACCGGAGTAAA------
     21                                                                                GAACGA---
```

Fig. 2. (Futami)

```
CAM1                                                                    -1374 GGATCCCGTGCCCCTGGTTG-AGAGACACTGTTC -1341
CAM2                                                                    -1537 ......A.G..T.......CA....-.....C.. -1504

CAM1 TGGATAAC-TTG---CATTTTGGTGACATGCTGAGGTGGTAGTTGAAGATGGTTACATTTCTTGATTGTTTTGTTGCAACAACATACTCC -1255
CAM2 .....G..T...ACTT.CA..T....T..........A...T...TT......................A..............TA... -1414

CAM1 ATAGGCTACATG---TCTGCATTCTTATAAACTTATTGATTTCTATGAACATGGTTTCTATAATCAATGCTGTCACTAGTTTTT------ -1174
CAM2 ....C.A.T...TAG.......C...............G......CA.G...........C.GC.C...T........CG.G..TATATA -1324

CAM1 -----------------------------------TTTTTATATGC-TTTGTCACAGTCATACAAATTCTAGAAACACTTTGTAGTCAG -1121
CAM2 TATATATATATATATATATATATATATATATATATA.A.A.....A.T.......A....T......................C.----.. -1238

CAM1 CACACAAAATATAAGAT--------------AAAAAAT--AAAC-----AAACAAA-CAAA-CAAACAGGGGTATAAGAAGATATCTAAT -1054
CAM2 .........G...T..CAAATTTTAGCACGC.......GA...ATGAAG...T...GG...TT...A.TACAA.G..........A..-G -1149

CAM1 TTCACATTAA-GAGAGTGCTTCTAGGCTTTG-------C----ACAATG--TTATTTCTATTTAAAACAACAAAAGGGTGATTTCTTTTA -978
CAM2 .......C..G.T.TT..G.G..TT......GTTTCCT.AGTA......TT.......A.T........T.......A....G....C.CC. -1059

CAM1 ACCTATTTTGCTTATCTTTGTGGTA-TTTAAACA---GAAT--------A-----------AGACAGAAATAAACAGGAACGGTCTGATT -911
CAM2 ........A....T.A.....T..TG...TTTTGTTT.TT.GTTTGTTT.TTGATTTTTTT.C.G.AT...TT.T.A...TTA.A..... -969

CAM1 TTGTTGTATCGTTGTTTACTAAAATGAAAATTCATCCTTTTTTCAAAAAA-GTGAGCCGTCA---------------------------- -850
CAM2 ......C...C.....A..GT...AAC.C.-----......---A......T......T....CTGCTCGACTATAAACGATCCTTCACTC -886

CAM1 CAATGACGTCAAATCCCCGCCCACCA--G-CTTATCG------------TTCCGCGAGAAAATAGTTCCACATTTATCGAATTCAAAAAT -775
CAM2 .......................T..TA.A...C.GCCAACATATGCTC..................C....G.............G.... -796
                            ↓
CAM1 AGATGTAACATGTCCTCTAAGTCACTAAAGACTAATTATAGTTTCAAAGACGTCGTTGTGTTTTTAGTTCTTATGTGTGAGCCTGGC--G -687
CAM2 .A...TA..A.....T....CA...A............AG.G.....CG..G.................T....C.............A..GA. -706

CAM1 TGACATTTAATACGCAAACAACGGTGTTTTCGTTTGGCGGGAAACGTTAGTGCGCAGTCATAAAGTCAGTGGGCGCTGCTTCAGCTCGCG -597
CAM2 G.GG.......T...........AA............................G.............-.C....-------A....G.T....←. -624

CAM1 GGAGCTACAGAACTTATTCATCCTGCCACTGACACCCAGGCTATCAGTCTGTCCAGCACCTGAGGGGTGCTCGAGTGCGTACGTCGATTA -507
CAM2 ............T........C......T...T...A..A.A...A.C........T..A....T.....A....A.ATAG.A....AG...T..... -534
       Myc-2B     ←    Myc-2A          ↑        ↓↑ ↑Inr ↓     ↑                              ↑
CAM1 CGGTTTCCATCACTTTCTATTTGGCCAGGACTCTGATTCCACTTATGCTGCAAAAGACCGGAGTAAACTATTGGATCTTACGTTTTGTTC -417
CAM2 .AT.....A........CC......A..................C..C......T..G...G...............G...........T......C.. -444
            ↑         ↑                     ↓  ↑  ↑        ↑
CAM1 CTTATTTTTTGCTCTCTTTTCATTGCATTGCGTCTCGAACGAAATCATTCTGAAACTGCATTTGTTAAA-----TACACGgtaaggatcc -332
CAM2 ..G.........T...........C.G...C.C.........GA...................................TAACT................a -354
                                 ↑     ↑
CAM1 gacatttgatgctgcatgcatcttttagcctt-----------------aatatgaatgcaaatatctttgt-tgttaaatgggagtttgc -259
CAM2 ...........t...................a...aaagtgcaaattgcgtgt.....g........g.a....a.a.t..t..aa.....a....a -264

CAM1 at-gtt-gcacgttcaaagaatacact-agaagcagggctgtttttaatg-ttttgatagcaaccccgctttgaggtttaatcgctccg -173
CAM2 ..g...g.....c................a.a......a.....c....a.....t.........t.............tg...........t..... -174

CAM1 tttttacacatactgtcatataacttagttcctttatttagctattttatccaccc-catgcatgttatgcatagtctaaatatgtgcat -84
CAM2 ....c.......g.....------....g......a...............g...tc.......t........acg............t...............- -90

CAM1 tttaggtaaagc-----gttgtgatgtttagttagacgttgcatatgcacgataataattccgatgtgta-tttttcacaagTGTCACA -1
CAM2 -......c.....gtttt.........a.g............................tg..g.........a..t.....t...................A. -1

CAM1 ATGCCGGTGAGTGCGAGTTTGGCATATAAAAACTACGATTACGACTACGACTCCATCCAGCCCTACTTCTACTTCGACAACGACGATGAA +90
CAM2 ATG.←......T...........G..........................T..................T..........................G +90
           Myc-1B
CAM1 GATTTCTATCACCATCAGCAAGGACAGACTCAACCTCCAGCTCCCAGCGAGGACATTTGGAAAAAATTCGAGCTGCTGCCCACACCGCCC +180
CAM2 ←............................................................G............................ +180
          Myc-1A
```

Fig. 2. (Futami)

第 3 章　コイ ***Max*** の cDNA クローニングと、c-***myc*** との発現の比較

Futami, K., Komiya, T., Zhang, H. and Okamoto, N.: Differential expression of *max* and two types of c-*myc* genes in a tetraploid fish, the common carp (*Cyprinus carpio*). *Gene*, **269**,113-119 (2001)

Differential expression of *max* and two types of c-*myc* genes in a tetraploid fish, the common carp (*Cyprinus carpio*)

Keywords: alternative splicing; cloning; evolution; insertion; polyploidy

Kunihiko Futami [a], Takeru Komiya [a], Huan Zhang [b], Nobuaki Okamoto [a,*]

[a] *Department of Aquatic Biosciences, Tokyo University of Fisheries, Konan 4, Minato-ku, Tokyo 108-8477, Japan*

[b] *Department of Marine Sciences, University of Connecticut, Groton, CT06340, USA*

*Corresponding author. Tel: +81 3 5463 0547, Fax: +81 3 5463 0552; e-mail: nokamoto@tokyo-u-fish.ac.jp

Abbreviations: aa, amino acid(s); bp, base pair(s); *CAM*1, carp c-*myc*1 encoding c-MYC1; *CAM*2, carp c-*myc*2 encoding c-MYC2; cDNA, DNA complementary to RNA; c-*myc*, gene(s) encoding c-MYC; c-MYC, cellular MYC; kb, kilobase(s), nt, nucleotide(s); PCR, polymerase chain reaction; *max*, *myc* associate protein X.

ABSTRACT

We cloned the full-length cDNA of *max* gene from the common carp (*Cyprinus carpio*). The cDNA clone of carp *max* consists of 1209 bp and contained an ATG-initiated ORF consisting of 156 aa. The carp MAX share 76.7-93.8 % aa identity with those of human, mouse, rat, chicken, *Xenopus* and zebrafish, respectively. The 15 bp alternative splicing was observed in the loop region of helix-loop-helix and is not previously described in mammalian *max* sequences. Transcripts of *max* gene were observed in all of the tissues of carp investigated in this study. The highest expression was found in the ovary, and the transcripts in hepatopancreas and heart were low. Two carp c-*myc* genes (*CAM*1 and *CAM*2) showed differential expression pattern. The expression of *max* was concomitant with *CAM*2 expression, but not with *CAM*1. It has been reported that MYC/MAX heterodimer as a regulator of gene expression has been maintained throughout vertebrate evolution, and the expression of c-*myc* has been concomitant with *max* expression. In addition, according to phylogenetic analysis, *CAM*1 is evolving faster than *CAM*2 after gene duplication. Therefore, this result suggests that *CAM*1 may evolve to obtain a new function different from c-*myc*.

Introduction

The proto-oncogene c-*myc* is thought to be one of the most important genes in controlling cell proliferation (Roy et al., 1993). It has precise expression (both specifically and quantitatively), is crucial for cell division and differentiation and is highly conserved in vertebrates. However, its mode of action and its interaction with the signaling pathway is still unclear. In mammals, c-*myc* genes consist of three exons and two introns (Bernard et al., 1983). The first exon is a noncoding exon. It plays a regulatory role in the transcription of the c-*myc* gene (Saito et al., 1983). Furthermore, in the human genome, 5 members of the *myc* gene family (c-*myc*, L-*myc*, N-*myc*, s-*myc*, B-*myc*) have been reported (Ryan and Birnie, 1996). Each member is structurally similar to one another, however, nucleotide and amino acid identities shared by them are less than 60%. Evolutionary origin and relationships of each *myc* member remain unknown.

MAX is a basic helix-loop-helix/leucine zipper (bHLH/LZ) protein, which forms heterodimers with members of the MYC protein family (Blackwood and Eisenman, 1991). MYC/MAX heterodimers exhibit sequence-specific DNA binding with much greater affinity than MYC homodimers. MAX may also form homodimers which recognize the same target sequence as the MYC/MAX heterodimer, but which are unable to function as transcription activators (Amati et al., 1992; Kato et al., 1992).

In lower vertebrates, both c-*myc* and *max* genes have already been isolated from the zebrafish, and it is suggested that MYC/MAX heterodimer as a regulator of gene expression has been maintained throughout vertebrate evolution (Schreiber-Agus et al., 1993a). In a tetraploid fish, two c-*myc* genes (*CAM*1 and *CAM*2) have been isolated

from the common carp, *Cyprinus carpio* (Zhang et al., 1995). According to phylogenetic analysis, *CAM*1 is evolving faster than *CAM*2 after gene duplication (Zhang et al., 1994). In addition, we determined the heterogeneous transcription start points of two c-*myc* genes from the carp as reported previously (Futami et al., 2000). The first exons of the carp c-*myc* genes are evolving faster than the second and third exons. The differences in exon1 and the promoter structure between the two c-*myc* genes of carp suggested that *CAM*1 and *CAM*2 were evolving to acquire different functions after the tetraploid event. However, *max* gene had not been isolated from the common carp, so relations between *max* and two c-*myc* genes had not been proven.

Polyploidy is a potentially important process in the evolution of vertebrates (Ohno, 1970; Lundin, 1993). Studies on gene duplication in tetraploid teleosts are important for investigating the evolutionary processes following the tetraploid event (Ohno, 1993). Furthermore, the study of expression of c-*myc* and *max* may help us to understand evolutionary origin of *myc* gene family in vertebrates, besides knowing the transcriptional function of *max* and two types of c-*myc*. In this report, we cloned a *max* gene from the common carp and compared it with expression patterns of two types of c-*myc* genes in carp adult tissues.

2. Materials and methods

2.1. Isolation of max cDNA

Due to the fact that *max* was detected in the uterus and developmental stage of zebrafish (Schreiber-Agus et al., 1993a) and carp, by preliminary experiment, the ovary was selected for total RNA extraction. Extraction was performed using TRIZOL regent (Gibco BRL), according to the manufacture's protocol. Two μg of total RNA was subjected to reverse transcription by reverse-transcriptase (SUPERSCRIPT II, Gibco BRL) using oligo-dT primer. Using a fiftieth of cDNA as a template, the coding region of *max* was amplified by PCR. The primers were P1 (5'-ATG AGC GAC AAC GAT GAT ATC GAG G-3') and P2 (5'-TCC TCC GGG CGA TGC TTC TT-3'), which were designed based on the reported sequence of the *max* gene of zebrafish. PCR amplification was started with a 2 min hold at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C with a post-extension of 5 min at 72°C. The amplified fragments were separated by an agarose gel electrophoresis, and cloned into a home-made T-tailed pBluescript II SK(-) vector. Sequences were analyzed by dye terminator cycle sequencing using the ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Cetus, USA).

The 5' upstream region of carp *max* was determined by the 5' RACE method (Frohman et al., 1988). 0.5 μg of total RNA was subjected to reverse transcription by reverse-transcriptase (SUPERSCRIPT II, Gibco BRL) using *max* specific primer P2. After hydrolysis of the RNA with RNase H, cDNA was purified with GENECLEAN II Kit (BIO 101), and subjected to oligo-dC tailing reaction with terminal

deoxynucleotidyl transferase (Gibco BRL). The PCR reaction of dc-tailed cDNA was performed using AmpliTaq Gold (Perkin-Elmer Cetus, USA) with 0.1 μM of PCR primers. The primers were anchor primer (Pa, 5'-GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG-3') and *max* specific primer 3 (P3, 5'-GCT GGT GTG TGT GGT TTT TCC GTC-3'). PCR Amplification was started with a 12 min hold at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 57°C, and 1 min at 72°C with a post-extension of 5 min at 72°C. The primary reaction products were used as the template for the secondary amplification of nested PCR. In this secondary reaction, universal amplification primer (Pb, 5'-GGC CAC GCG TCG ACT AGT AC-3') and *max* specific primer (P4, 5'-GCT GTC TTT GAT GTG GTC CCT ACG-3') were used. PCR was performed same as the first PCR. The PCR fragment was subcloned into pBluescript II SK(-) and sequenced.

3' franking region was determined by the 3' RACE method. Two μg of total RNA was subjected to reverse transcription by reverse-transcriptase (SUPERSCRIPT II, Gibco BRL) using oligo(dT)-containing adapter primer. The PCR reaction was performed using AmpliTaq Gold (Perkin-Elmer Cetus, USA) with 0.1 μM of PCR primers. The primers were adapter primer (Pc, 5'-GGC CAC GCG TCG ACT AGT AC-3') and *max* specific primer P1. PCR amplification was started with a 12 min hold at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 58°C, and 1.5 min at 72°C with a post-extension of 5 min at 72°C. The amplified fragments were separated by an agarose gel electrophoresis. The predicted 1-1.2 kb products were eluted from the gel and used as the template for the secondary amplification of nested PCR. The primers were Pc and *max* specific primer (P5, 5'-CGG AAA AAC CAC ACA CAC CAG

CAG-3'). PCR was performed the same as the first PCR. The PCR fragment was subcloned into pBluescript II SK(-) and sequenced.

2.2. Genomic Southern blot analysis

10 μg of carp genomic DNA was digested completely with *Eco*RI or *Hin*dIII, and electrophoresed in 0.8 % agarose gel and transferred with 0.4 N NaOH to a nylon membrane (Hybond N+, Amersham Falmacia Biotech). The blot was hybridized with the ^{32}P-labeled probe. The probe used was a part of putative exon1 of carp *max*, which was amplified by PCR using cDNA clone as template. Membrane hybridization as well as washing procedures were carried out at 65°C, according to the standard protocol (Sambrook et al., 1989).

2.3. Northern blot analysis

To analyze the tissue-specific expression of the *max* gene, total RNA was extracted from hepatopancreas, kidney, brain, heart, gill and ovary of adult carp. Twenty μg of total RNA was electrophoretically separated on 1% agarose/formaldehyde gels and blotted onto a nylon membrane (Hybond N+, Amersham Falmacia Biotech). The blot was hybridized with the ^{32}P-labeled cDNA probe. Membrane hybridization as well as washing procedures were carried out at 65°C, according to the method developed by Church and Gilbert (1984).

2.4. Detection of carp two c-myc mRNA by RT-PCR/Southern blot hybridization

Two μg of total RNA was subjected to reverse transcription by M-MLV reverse transcriptase (Promega) using oligo-dT primer. A fiftieth of cDNA was used for a PCR reaction. The primer set P6-P7 was used for RT-PCR of *CAM*1 and P6-P8 for *CAM*2 (P6: 5'-GCT TT CCG CTG CTG CCA AGT T-3'; P7: 5'-GTA CCT TGA ATC TGA CAC TGC CGT-3'; P8: 5'-TAC CTT GAA TCG GAC ACC TCT GC-3'). The specificity of the PCR was confirmed by sequencing. Expression of cytoskeletal β-actin gene was used for internal control. The primers for β-actin gene were designed based on the reported sequence (Katagiri et al., 1997). PCR Amplification was started with a 2 min hold at 95°C, followed by 20 cycles of 30 s at 94°C, 30 s at 56°C, and 30 s at 72°C with a post-extension of 3 min at 72°C. The PCR products were not detected by ethidium bromide staining, so the reaction products were electrophoresed on a 2% agarose gel and then transferred to a nylon membrane. The PCR products were then detected by Southern blot hybridization (Sambrook et al., 1989).

3. Results and discussion

3.1. Isolation of carp max cDNA clone

An RT-PCR strategy for cloning a partial cDNA of carp *max* successfully yielded a cDNA fragment of 419 bp. Database searches with the partial nt sequence thereof using BLAST program (Altschul et al., 1990) invariably yielded high scores of similarity to other vertebrate *max* sequences, and the most closely related to zebrafish *max* (data not shown). In order to obtain a full-length cDNA of carp *max*, the 5' and 3' RACE were performed (Fig. 1). As a result, the nt sequences of carp *max* cDNA clone, 1209 bp in length, were determined (DDBJ, registration number AB036771). This nt sequence contained an ATG-initiated ORF consisting of 156 aa (Fig. 2). Using "GENETYX-MAC" computer algorithm developed by Software Development Co., the carp MAX share 82.3 %, 76.7 %, 76.7 %, 77.4 %, 78.2 % and 93.8 % aa identity with those of human, mouse, rat, chicken, *Xenopus* and zebrafish, respectively. Notably, the bHLH/LZ region (Blackwood and Eisenman, 1991) was highly conserved throughout vertebrate evolution (Fig. 3).

The mammalian *max* gene has been shown to encode several alternatively processed transcripts (Blackwood and Eisenman, 1991; Prendergast et al., 1991). Carp *max* transcripts also undergo 15 bp alternative splicing. However, this alternatively spliced sequence is not previously described in mammalian *max* sequences, and only fish *max* genes contain this insertion, although there is no comment that this inserted sequence is alternatively spliced in zebrafish. Furthermore, this alternative splicing is observed in the loop region of helix-loop-helix. Therefore, a *max* cDNA of carp

encodes two members of MAX isoforms, which may bind some proteins. However, the differential activities of these alternative forms remain undetermined in this report.

Two bands were observed by genomic Southern blot analysis, in *Eco*RI digests (21 kb and 4.6 kb) and *Hin*dIII digest (4.0 kb and 2.4 kb), respectively (Fig. 4). These bands are thought to correspond to at least two *max* genes existing per haploidy genome in the common carp, because carp are tetraploidy. Although we isolated sixteen single clones from a RT-PCR product and analyzed the nt sequences, other *max* clones were not isolated. Therefore, after the tetraploid event, one of the 2 duplicated genes may not be transcribed in the tissue examined in this study or become a pseudogene.

3.2. Differential expression of max and two c-myc genes in several tissues

Transcripts of *max* gene were observed in all of the carp tissues (hepatopancreas, kidney, brain, heart, gill, ovary) investigated in this study (Fig. 5). The highest expression was found in the ovary, and the transcripts in hepatopancreas and heart were low. Dramatically expression in ovary may correlate with L-*myc* which is transcribed in the uterus and during early development (Schreiber-Agus et al., 1993a).

We also analyzed the tissue-specific expression of two c-*myc* genes of carp by RT-PCR/Southern blot hybridization (Fig. 6). In mammalian terminally differentiated tissues, c-*myc* expression is low or absent altogether (Zimmerman et al., 1986). In contrast, transcripts of two c-*myc* genes were observed in all of the tissues investigated in this study. In addition, high expression of c-*myc* in various adult tissues of zebrafish and *Xenopus* was also observed (Schreiber-Agus et al., 1993a; Schreiber-Agus et al., 1993b). The steady-state expression in lower vertebrates may reflect with lower

vertebrate-specific functions, such as tissue regeneration and/or immortalization of cell lines. However, two carp c-*myc* mRNAs were not clearly detectable by Northern blot hybridization and RNase protection assay using ten μg of total RNA in any of the organs investigated (data not shown). This result suggests that the level of c-*myc* expression in carp tissues examined in this study may be low.

Comparing *CAM*1 with *CAM*2, mRNA level of *CAM*2 in the hepatopancreas was lower than that of *CAM*1, while in the ovary, mRNA level of *CAM*2 was higher than that of *CAM*1. The differences of expression pattern between the two c-*myc* genes of carp suggested that *CAM*1 and *CAM*2 were evolving to acquire different functions after the tetraploid event. In our previous study, we determined the heterogeneous transcription start points of two c-*myc* genes from the hepatopancreas of carp (Futami et al., 2000). Differential expression pattern of two c-*myc* genes in the hepatopancreas may correlate with the variations of *tsp*.

Interestingly, although the expression pattern of carp *max* is similar to that of *CAM*2, it is not similar to that of *CAM*1. In zebrafish, it is suggested that MYC/MAX heterodimer as a regulator of gene expression has been maintained throughout vertebrate evolution, and the expression of c-*myc* has been concomitant with *max* expression (Schreiber-Agus et al., 1993a). Therefore, the coordinate expression of *CAM*2 in the same tissues suggests that the CAM2/MAX complex may serve an active physiological role as an original MYC/MAX heterodimer. In contrast, *CAM*1 may evolve to obtain a new function different from c-*myc*. Indeed, according to phylogenetic analysis, *CAM*1 is evolving 1.6 times faster than *CAM*2 after gene duplication, and *CAM*2 is conserved throughout vertebrate evolution. (Zhang, 1994). This result agrees with the suggestion that subsequent to the tetraploidization event, one

of the 2 duplicated genes may evolve faster to obtain a new function or become silent (Ohno, 1970). However, the MOTIF program (http://motif.genome.ad.jp/) to predict the 3D structures of MAX, CAM1 and CAM2 showed that all of these had the motif of helix-loop-helix (data not shown). Therefore, the CAM1/MAX heterodimer formation may be considered as a possibility. It remains to be analyzed about the difference of affinity *in vivo* between MAX and two c-MYC.

The present study may help us to understand the evolutionary origin and relationships of the *myc* gene family in vertebrates, besides knowing the transcriptional control and evolution of *max* and two types of c-*myc* genes in tetraploid fishes. Further studies are needed to determine the difference of intracellular function between two c-*myc* genes, and the protein-protein interaction of MAX and two c-MYC.

3.3. Conclusions

(1) We cloned the carp *max* cDNA by the RT-PCR and the RACE method. The aa sequence of this gene were highly conserved throughout vertebrate evolution.

(2) Carp *max* transcripts undergo 15 bp alternative splicing. This alternatively spliced sequence is not previously described in mammalian *max* sequences, and only fish *max* genes contained this insertion

(3) Transcripts of the *max* gene were observed in all of the carp tissues (hepatopancreas, kidney, brain, heart, gill, ovary) investigated in this study. The highest expression was found in the ovary. Two carp c-*myc* genes (*CAM*1 and *CAM*2) showed differential expression pattern. The expression of *max* was concomitant with

CAM2 expression, but not concomitant with *CAM1*. This result suggests that *CAM1* may evolve to obtain a new function different from c-*myc*.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E., Lipman, D. J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Amati, B., Dalton, S., Brooks, M. W., Littlewood, T. D., Evan, G. I., Land, H., 1992. Transcriptional activation by the human c-Myc oncoprotein in yeast requires interaction with Max.. Nature 359, 423-429

Bernard, O., Cory, S., Gerondakis, S., Webb, E., Adams, J. M., 1983. Sequence of the murine and human cellular *myc* oncogenes and two modes of *myc* transcription resulting from chromosome translocation in B lymphoid tumors. EMBL J. 2, 2375-2383.

Blackwood, E. M., Eisenman R. N., 1991. Max: A Helix-Loop-Helix Zipper Protein That Forms a Sequence-Specific DNA-Binding Complex with Myc. Science 251, 1211-1217.

Church, G. M., Gilbert, W., 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA 81, 1991-1995.

Frohman, M. A., Dush, M. K., Martin, G. R. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. U S A 85, 8998-9002.

Futami, K., Komiya, T., Zhang, H. Okamoto, N. 2000. Determination of heterogeneous transcription start points of two c-*myc* genes from the common carp (*Cyprinus carpio*). Gene 245, 43-47.

Katagiri, T., Hirono, I., Aoki, T., 1997. Identification of a cDNA for medaka cytoskeletal â-actin and construction for the reverse transcriptase- polymerase chain reaction (RT-PCR) primers. Fisheries Science 63, 73-76.

Kato, G. J., Lee, W. M., Chen, L. L., Dang, C. V., 1992. Max: functional domains and interaction with c-Myc. Genes. Dev. 6, 81-92.

Lundin, L. G., 1993. Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16, 1-19.

Ohno, S. (1970). Evolution by gene duplication. Springer Verlag Press, Heidelberg, New York.

Ohno, S., 1993. Patterns in genome evolution. Current Opinion in Genetics and Development 3, 911-914.

Prendergast, G. C., Lawe, D., Ziff, E. B. 1991. Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell 65, 395-407

Ryan, K. M., Birnie, G. D., 1996. *Myc* oncogenes: the enigmatic family. Biochem. J. 314, 713-721.

Roy, A. L., Carruthers, C., Gutjahr, T., Roeder, R. G., 1993. Direct role for Myc in transcription initiation mediated by interactions with TF II-I. Nature 365, 359-361.

Saito, H., Hayday A. C., Wiman, K., Hayward, W. S., Tonegawa, S., 1983. Activation of the c-*myc* gene by transcription: A model for translation control. Proc. Natl. Acad. Sci. USA 80, 7476-7480.

Sambrook, J., Fritsch, E. F., Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Schreiber-Agus, N., Horner, J., Torres, R., Chiu, F. C., Depinho, R. A., 1993. Zebra Fish *myc* Family and *max* Genes: Differential Expression and Oncogenic Activity throughout Vertebrate Evolution. Mol. Cell. Biol. 13, 2765-2775.

Schreiber-Agus N., Torres R., Horner J., Lau, A., Jamrich, M., Depinho R. A., 1993. Comparative Analysis of the Expression and Oncogenic Activities of *Xenopus* c-, M-, and L-*myc* Homologs. Mol. Cell. Biol. 13, 2456-2468.

Thompson, J. D., Higgins, D. G., Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic. Acids Res. 22, 4673-4680.

Zhang, H., Okamoto, N., Ikeda, Y., 1995. Two c-*myc* genes from a tetraploid fish, the common carp (*Cyprinus carpio*). Gene 153, 231-236.

Zhang, H., 1994. Isolation and DNA Sequence of one c-*myc* Gene from Goldfish *Carassius auratus* and the Molecular Phylogeny for c-Myc of Vertebrates. In: Doctoral dissertation, Tokyo University of Fisheries.

Zimmerman, K. A., Yancopoulos, G. D., Collum, R. G., Smith, R. K., Kohl, N. E., Denis, K. A., Nau, M. M., Witte, O. N., Toran-Allerand, D., Gee, C. E., Minna, J.D., Alt, F. W., 1986. Differential expression of *myc* family genes during murine development. Nature 319, 780-783.

Figure legend

Fig. 1. Strategy of cloning the full-length cDNA of *max* gene of common carp. The full-length cDNA of *max* was determined by 3 overlapped partial cDNA, 5'RACE, RT-PCR, 3'RACE products. Pa-Pc, P1-P5, primers; A, adapter; 5'UTR, 5' untranslated region; ORF, open reading frame; 3'UTR, 3' untranslated region.

Fig. 2. The nt and deduced aa sequences of carp *max* cDNA. Deduced amino acids are shown as one letter code below each codon. Alternatively spliced sequence is underlined

Fig. 3. Alignment of deduced amino acid sequences of the *max* genes of chicken (EMBL, L12469), *Xenopus laevis* (L09738), rat (D14447), mouse (M63903), zebra fish (L11711), common carp (DDBJ, AB036771) and human (EMBL, M64240),. The abbreviated standard one-letter code aa sequences were initially aligned by using a multiple alignment program in CLUSTAL W (Thompson et al., 1994). Asterisks (*) represent identity between the seven animals. Gaps (-) were introduced to optimize identity. Alternatively spliced sequence is underlined. The bHLH was indicated. Conserved hydrophobic residues of the leucine zipper were indicated by sharps (#).

Fig. 4. Genomic Southern blot analysis of carp genomic DNA (10 μg) digested with *Eco*RI (*lane 1*) or *Hin*dIII (*lane 2*) and hybridized with a part of putative exon1 of carp *max*, which was amplified by PCR using cDNA clone as template. Two bands were detected in *Eco*RI digests (21 kb and 4.6 kb) and *Hin*dIII digest (4.0 kb and 2.4 kb), respectively

Fig. 5. Tissues distribution of *max* mRNA expression in carp by Northern blot analysis.

Fig. 6. Detection of the mRNA of carp two c-*myc* genes by RT-PCR/Southern hybridization. Expression of cytoskeletal β-actin gene was used for internal control. The data shown were derived from a single experiment that is representative of at least two independent experiments.

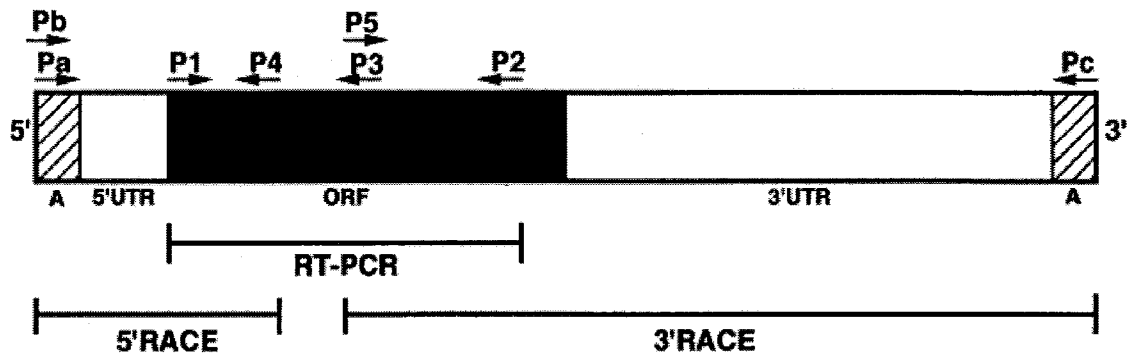


Fig. 1. (Futami)

```
 -97   ACAGGCGTGACACGCCACGGTAAATGTCCGTGTGTAGGCGAGTCACGTGTGGAGAGTCCG   -38

 -37   TTGATCTCGTTTCGTATTTGTCTATCTTCGCCGAAGAATGAGCGACAACGATGATATCGA   23
                                                     M  S  D  N  D  D  I  E

  24   GGTCGACAGTGATGCAGACAAACGGGCACATCACAATGCGCTGGAGCGCAAACGTAGGGA   83
         V  D  S  D  A  D  K  R  A  H  H  N  A  L  E  R  K  R  R  D

  84   CCACATCAAAGACAGCTTTCACAGCCTTCGGGATTCCGTTCCCGCCTTGCAAGGGGAAAA   143
         H  I  K  D  S  F  H  S  L  R  D  S  V  P  A  L  Q  G  E  K

 144   GCAATCTATCAAACAGGCATCCCGAGCTCAAATCCTAGACAAAGCCACAGAGTACATCCA   203
         Q  S  I  K  Q  A  S  R  A  Q  I  L  D  K  A  T  E  Y  I  Q

 204   GTACATGCGACGGAAAAACCACACACACCAGCAGGACATCGACGACCTGAAGAGGCAGAA   263
         Y  M  R  R  K  N  H  T  H  Q  Q  D  I  D  D  L  K  R  Q  N

 264   CGCTCTGCTGGAGCAACAAGTACGGGCACTGGAGAAAGTCAATGGGACCACGCAGCTGCA   323
         A  L  L  E  Q  Q  V  R  A  L  E  K  V  N  G  T  T  Q  L  Q

 324   GGCCAACTACTCCTCTTCAGACAGCAGCTTGTACACCAACCCCAAGGGCAGCGCTGTATC   383
         A  N  Y  S  S  S  D  S  S  L  Y  T  N  P  K  G  S  A  V  S

 384   GGCCTTCGACGGTGGTTCCGACTCGAGCTCAGAGTCTGAGCCGGAGGAACAGCGTTCCCG   443
         A  F  D  G  G  S  D  S  S  S  E  S  E  P  E  E  Q  R  S  R

 444   GAAGAAGCCCCGTGGGGAGGACAGCTAAACAGCGAATCCTCTGCCTGCTCCACAAAGTTC   503
         K  K  P  R  G  E  D  S  *

 504   TCCTCTGAGGGGAGGGAGAGTTACAGCAATGGTCTTGTTTCTTCTTGTTTTTTGTGTACC   563

 564   TCAATCGCCCCAAACCACCTTGATCAATCTCTTTCAGTGAATGTCAACGCTCGATTCCAC   623

 624   GCCTGAAGAAACCTCTTCCAGGTTTATAAGAGAGAGGGACACCTGTTCCGGGAAGAACAC   683

 684   TCACGGAAAACGGCTTGTTTTTACTCTCCCTCCCTCTCCGACTCTCTTCGTCCATCTTGT   743

 744   GTCCTGGCACACGCAACAACAAATGATGATTTAAACCCAAAATGAGGCAGCTTTGCAACT   803

 804   TAAGGACTTGATGCTTTGTACCCTTGTCTGCAATCCCTCCAGGTGGATATCCAAAGGGTT   863

 864   GAATGAGGCGTACAGCTCCTAGCACTGCTAAAATATACTTTTTTTGTTGATTATCTTAAA   923

 924   GCCTGCTTAGAACTTACCCTTAGCTGATTTGTCCCGAGTGGTTTGCCTTTCTAAATATTT   983

 984   GTTACTTTTCCATTAAGGAATGTTTTTGAAGCATCTATCAATGTACCTCTTGTTAGACCT   1043

1044   AGGGGATGAGTGATACGTTTTGACGTTATCTATAAAATCCCTTATTTAAAAAATACAAAA   1103

1104   AAAAAAAAA   1112
```

Fig. 2. (Futami)

```
                                         basic region      Helix I          Loop
Chicken      MSDNDDIEVESDEEQPRFQSAADKRAHHNALERKRRDHIKDSFHSFTLRDSVPSLQ--GE
Xenopus      MSDNDDIEVESDEDSSRFPYSADKRAHHNALERKRRDHIKDSFHGFTLRDSVPALQ--GE
Rat          MSDNDDIEVESDEEQPRFQSAADKRAHHNALERKRRDHIKDSFHSFTLRDSVPSLQ--GE
Mouse        MSDNDDIEVESDEEQARFQSAADKRAHHNALERKRRDHIKDSFHSFTLRDSVPSLQ--GE
Zebrafish    MSDNDDIEVDSD---------ADKRAHHNALERKRRDHIKDSFH--SLRDSVPALQFTGE
Carp         MSDNDDIEVDSD---------ADKRAHHNALERKRRDHIKDSFH--SLRDSVPALQ--GE
Human        MSDNDDIEVESD---------ADKRAHHNALERKRRDHIKDSFH--SLRDSVPSLQFTGE
             *********:**         ***********************  :******:**  **

                          Helix II
Chicken      K-----ASRAQILDKATEYIQYMRRKNHTHQQDIDDLKRQNALLEQQVRALEKFTARSSA
Xenopus      K-----ASRAQILDKATEYIQYMRRKNHTHQQDIDDLKRQNALLEQQVRALEKFTAKSSS
Rat          K-----ASRAQILDKATEYIQYMRRKNHTHQQDIDDLKRQNALLEQQVRALEKFTARSSA
Mouse        K-----ASRAQILDKATEYIQYMRRKNDTHQQDIDDLKRQNALLEQQVRALEKFTARSSA
Zebrafish    KQSIKQASRAQILDKATEYIQYMRRKNHTHQQDIDDLKRQNALLEQQVRALEKVKGTFTT
Carp         KQSIKQASRAQILDKATEYIQYMRRKNHTHQQDIDDLKRQNALLEQQVRALEKVNG--TT
Human        K-----ASRAQILDKATEYIQYMRRKNHTHQQDIDDLKRQNALLEQQVRALEKARSSAQL
             *     *********************.*************************     .
                            #      #      #      #      #      #

Chicken      QLQANYPAADSSLYTNPKGSTISAFDGGSDSSSDSEPDEPQSRKKLRMEAS
Xenopus      QLQSNY---------------------SSSESETEEPQSRKKLRMDAS
Rat          QLQTNYPSSDNSLYTNAKGGTISAFDGGSDSSSESEPEEPQNRKKLRMEAS
Mouse        QLQTNYPSSDNSLYTNAKGGTISAFDGGSDSSSESEPEEPQSRKKLRMEAS
Zebrafish    QLQANYSSSDSSLYTNPKGQAVSAFDGGSDSSSGSEPEEQRTRKKHRPEDS
Carp         QLQANYSSSDSSLYTNPKGSAVSAFDGGSDSSSESEPEEQRSRKKPRGEDS
Human        QTFTNYPSSDNSLYTNAKGSTISAFDGGSDSSSESEPEEPQSRKKLRMEAS
             *  :**                     *** **.:* :.*** * : *
```

Fig. 3. (Futami)

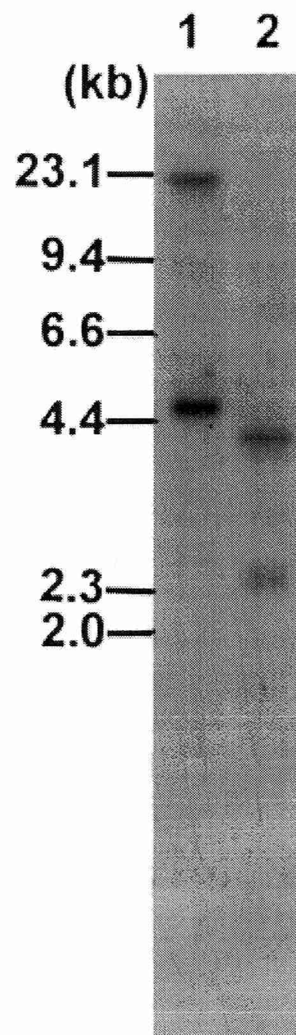


Fig. 4. (Futami)

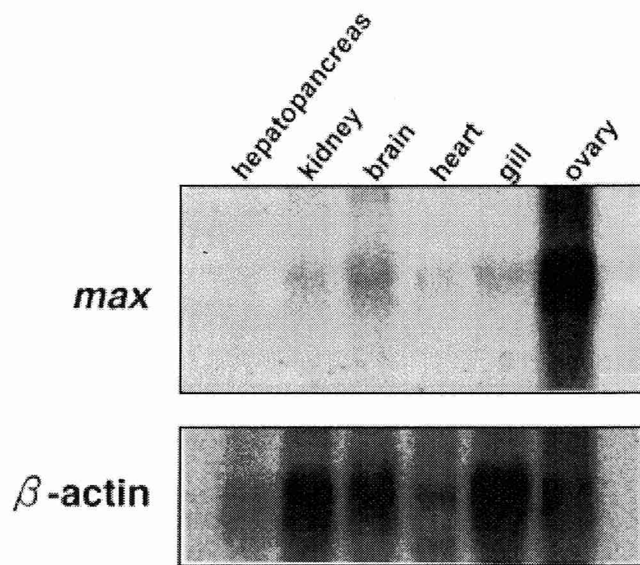


Fig. 5. (Futami)

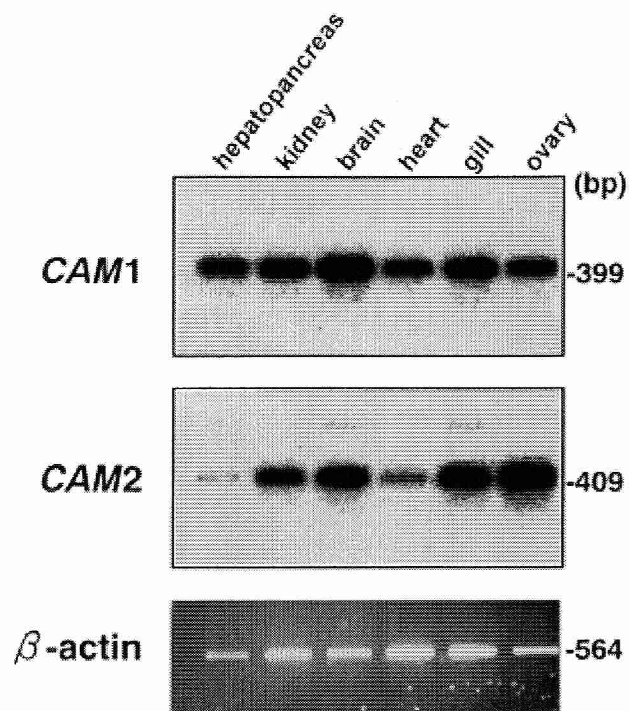


Fig. 6. (Futami)

第４章　４倍性魚類コイの c-*myc* 遺伝子２タイプの機能分化

Functional divergence of duplicated c-*myc* genes in a tetraploid fish, the common carp (*Cyprinus carpio*)*

Kunihiko Futami‡, Huan Zhang§, Nobuaki Okamoto‡¶

From the ‡Department of Aquatic Biosciences, Tokyo University of Fisheries, Konan 4, Minato-ku, Tokyo 108-8477, Japan and the §Department of Marine Sciences, University of Connecticut, Groton, CT06340, USA

*This work was supported in part by Research Fellowships of the Japan Society for the Promotion of Science (JSPS) for Young Scientists.

¶ To whom Correspondence should be addressed. Phone: 81-3-5463-0547; Fax: 81-3-5463-0552; E-mail: nokamoto@tokyo-u-fish.ac.jp

The abbreviations used are: bp, base pair(s); *CAM*1, carp c-*myc*1 encoding c-Myc1; *CAM*2, carp c-*myc*2 encoding c-Myc2; cDNA, DNA complementary to RNA; c-*myc*, gene(s) encoding c-Myc; c-Myc, cellular Myc; GST, glutathione S-transferase; kb, kilobase(s); PCR, polymerase chain reaction; *Max*, *Myc* associate protein X; PAGE, polyacrylamide gel electrophoresis.

Running Title: Functions of carp c-*myc* genes

SUMMERY

The proto-oncogene c-*myc* is thought to be one of the most important genes in controlling cell proliferation. In a tetraploid fish, two c-*myc* genes (*CAM*1 and *CAM*2) have been isolated from the common carp, *Cyprinus carpio*. Two carp c-*myc* genes (*CAM*1 and *CAM*2) showed different expression patterns in adult tissues as reported previously. Here we found that *CAM*1 and *CAM*2 expressed different patterns in cultured cells due to serum stimulation, and both *CAM*1 and *CAM*2 had distinct properties in terms of their transcription regulation system, cooperation with Max, and transcriptional activation to the *TERT* gene. These results showed that the two carp c-Myc have overlapping but distinct functions, suggesting that CAM1 and CAM2 may be evolving to acquire different functions after the tetraploidization event.

INTRODUCTION

Polyploidy is a potentially important process in the evolution of vertebrates (1, 2). It is believed that a tetraploid event took place about 500 million years ago in a common ancestor of all vertebrates (3). Higher vertebrates have evolved sex determination based on heteromorphic sex chromosomes, which prevent successful polyploidization (1). Polyploidy is still common in teleosts, which seem to be sexually undifferentiated. Studies on gene duplication in tetraploid teleosts are important for investigating the evolutionary processes following the tetraploid event (4).

The proto-oncogene c-*myc* is thought to be one of the most important genes in controlling cell proliferation (5). It is crucial for cell cycle, cell growth, differentiation, apoptosis, transformation, genomic instability and angiogenesis, and is highly conserved in vertebrates (6). Although a great number of researches have been carried out, the mode of action and its interaction with the signaling pathway is still unclear; the regulation of its expression and function is still far from being understood (7). In the human genome, 5 members of the *myc* gene family (c-*myc*, L-*myc*, N-*myc*, s-*myc*, B-*myc*) have been reported (8). Each member is structurally similar to one another, however, nucleotide and amino acid identities shared by them are less than 60%. Evolutionary origin and relationships of each *myc* member remain unknown.

In a tetraploid fish, two c-*myc* genes (*CAM*1 and *CAM*2) have been isolated from the common carp, *Cyprinus carpio* (9). According to phylogenetic analysis, *CAM*1 is evolving faster than *CAM*2 after gene duplication (10). In addition, *CAM*1 and *CAM*2 showed the heterogeneous transcription start points and different expression patterns as reported previously (11, 12). The expression of *Max*, which forms heterodimers with members of the Myc protein family (13), was concomitant with *CAM*2 expression, but

not concomitant with *CAM*1. In zebrafish, it is suggested that Myc/Max heterodimer as a regulator of gene expression has been maintained throughout vertebrate evolution, and the expression of c-*myc* has been concomitant with *Max* expression (14). Both the sequences and the pattern of expression lead us to predict functional differences. However, the biological significance of c-*myc*-duplication is unknown.

For better understanding of the evolution and function of *myc* family in higher vertebrates, studies on the expression and function of c-*myc* in lower vertebrates, especially in tetraploid fish, will provide useful insight. The biological significance of the duplication of c-*myc* is of great interest. Biochemical analyses at the protein level are required to reveal any functional divergence of the two c-*myc*. In this report, in order to clarify whether differentiation of some kinds of function occurs during two types of c-*myc* genes, we analyzed the differences of intracellular function between two c-*myc* genes, and the protein-protein interaction of Max and two c-Myc.

EXPERIMENTAL PROCEDURES

Cell culture

Three cell lines established from common carp were used in this study, KF (Koi Fin) and KG (Koi Gill) cells kindly provided by Dr. P. Hedrick, and EPC (Epithelioma Papulosum Cyprini) cells (15) were maintained in minimum essential medium (MEM) containing 10% fetal bovine serum and 14 mM N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid] (HEPES) at 20 °C

Plasmid construction

5' deletion mutants of *CAM*1 and *CAM*2 shown in Fig. 2 were made using Kilo-Sequence Deletion Kit (TAKARA BIO INC., Otsu, Japan) or PCR, and inserted in *Sma*I site of the pGL3-Basic vector (Promega). The cDNAs encoding both *CAM*1 and *CAM*2 were amplified by PCR-based strategy. The PCR products, which were produced by an additional *Bam*HI site in 5'-upstream and *Eco*RI site in 3'-downstream, were subcloned into the identical restriction sites of the expression vector pcDNA3.1(+) (Invitrogen) and pBIND (Promega). The pGL3-MMBS-SV40 plasmid was constructed by cloning into the *Mlu*I site of the pGL3-Promoter vector (Promega) the palindromic self-annealed synthetic double-stranded oligodeoxynucleotide 5'-CGC GGG AAG CAG ACC ACG TGG TCT GCT TCC-3', which includes the Myc/Max binding site (E-box, CACGTG) flanked by the *Mlu*I site (underlined). Full-length and alternatively spliced *Max* were inserted between the *Bam*HI and *Eco*RI sites of the pACT (Promega) vector. Plasmids for expression of GST fusion protein in *E. coli* BL21 were constructed by inserting full-length and alternatively spliced *Max* between *Eco*RI and *Bam*HI sites of the pGEX-3X (Amersham Bioscience).

Serum stimulation and semiquantitative RT-PCR

KF or KG cells were seeded in 24-well plates and cultured for 72 h in the growth medium containing 0.2% fetal bovine serum. Then the culture fluid were changed to growth medium containing 10% fetal bovine serum, and incubated for the indicated time.

Total RNA was isolated by using Sepasol RNA I Super (Nacalai Tesque), according to the manufacture's protocol. 0.5 (KF) or 0.1 (KG) μg of total RNA was subjected to reverse transcription by M-MLV reverse transcriptase (Promega) using oligo-dT primer. A fiftieth of cDNA was used for a PCR reaction. The primers used for RT-PCR of *CAM*1 and *CAM*2 were described before (12). The RT-PCR exponential phase was determined from 20 to 35 cycles to allow semiquantitative comparisons among cDNAs developed from identical reactions. All reactions involved an initial denaturation at 95 °C for 2 min followed by 25-32 cycles at 94 °C for 30 s, 56 °C 30 s and 72 °C for 30 s, on a Gene Amp PCR system 9600 (Perkin Elmer). Expression of cytoskeletal β-actin gene was used for internal control (16). The RT-PCR products were stained with SYBR Green I (Molecular Probes) and quantified by using Densitograph (ATTO).

Cloning of 5' flanking regions of carp c-myc *genes by inverse PCR*

To make template DNA, carp genomic DNA was digested with *Eco*RI, and self-ligated by T4 DNA ligase. Using about 100 ng of template DNA, we performed the first PCR amplification with the primer sets P1-P2 for *CAM*1 (P1: 5'-AAA TCC CCG CCC ACC AGC TTA TCG-3'; P2: 5'-TTC AAC TAC CAC CTC AGC ATG

TCA CC-3') and P3-P4 for *CAM2* (P3: 5'-TCA AAT CCC CGC CCA TCA TAG ACT TC-3'; P4: 5'-ACC ATC AAC AAA TAC TAC CTC AGC-3'). PCR Amplification was started with a 2 min hold at 95°C, followed by 35 cycles of 15 s at 95°C, and 4 min at 67°C with a post-extension of 3 min at 72°C. The primary reaction products of *CAM1* were used as the template for the secondary amplification of nested PCR. In this secondary reaction, nested primers P5 (5'-TCC GCG AGA AAA TAG TYC CAC RTT-3') and P2 were used. PCR was performed the same as the first PCR, but followed by 25 cycles. The PCR fragment was subcloned into home-made T-tailed pBluescript II SK(-) and sequenced by dye terminator cycle sequencing using the ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Cetus, USA).

Transfection and Luciferase assay

KF and KG cells were transfected with 1 μg various promoter mutants and 20 ng of pRL-SV40 vector using TransIT-LT1 (PanVera). After transfection, cells were incubated for 48 h. The luciferase activity was assayed by Pica-Gene Dual SeaPansy kit (Toyo Ink) and measured using GENE LIGHT 55 (Microtech Niti-On, Chiba, Japan). Relative luciferase activities were normalized by co-expressed *Renilla* luciferase in pRL-SV40 vector

Analysis of the methylation of CpG islands of c-myc *genes*

Three μg of KF and KG genomic DNA was digested completely with *Bsa*HI, which is methylation-sensitive restriction endonuclease, and electrophoresed in 0.8 % agarose gel and transferred to a nylon membrane (Hybond N+, Amersham Bioscience). The blot was hybridized with the ^{32}P-labeled probe. The probes used were 5' flanking

regions of *CAM*1 and *CAM*2. Hybridization procedures were performed using ULTRAhyb (Ambion), following the manufacturer's protocols.

GST-pull down assay

GST fusion proteins were prepared as described (17) except that the induction with isopropyl thio-D-galactoside was done at 30 °C. To synthesize the c-Myc proteins *in vitro*, TNT Quick Coupled Transcription/Translation Systems (Promega) was used. pcDNA3-*CAM*1 and pcDNA3-*CAM*2 was transcribed with T7 RNA polymerase and translated in the presence of [^{35}S]methionine. GST fusion proteins bound to GST beads (50 μl of 50% slurry), and 10 μl of *in vitro* translation products were mixed with 300 μl of NETN buffer (20 mM Tris, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40), incubated at on ice for 1 h, washed with NETN buffer 4 times, and subjected to SDS-PAGE (10% polyacrylamide gel) and fluorography.

Two-hybrid assay on fish cultured cells

Two-hybrid assay was carried out using CheckMate Mammalian Two-hybrid System (Promega) according to the following protocol. EPC cells were transfected at 20-30% confluence in 24-well dishes by using TransIT-LT1 (PanVera) following the manufacturer's instructions. 330 ng of pG5*luc* vector was cotransfected with 330 ng of pBIND-*CAM*1 or pBIND-*CAM*2, 330 ng of pACT-Max1 or pACT-Max2. In all assays, *Renilla* luciferase in pBIND vector was used as the internal control. After transfection, cells were incubated for 48 h. Cell extract preparations and dual luciferase assays were performed following the manufacturer's protocols (Toyo Ink).

Gel sift asssay

c-*myc* and *Max* transcripts were translated *in vitro* with nonradioactive methionine. c-Myc and Max (2:1) were mixed after translation, and analyzed for binding to the synthetic oligonucleotide containing the E-box (Myc/Max: 5'-GGA AGC AGA CCA CGT GGT CTG CTT CC-3') by the gel sift assay. For competition assay, mutant oligonucleotides for Myc/Max (5'-GGA AGC AGA CCA CGG AGT CTG CTT CC-3') were used as competitor. The DNA binding reaction was allowed to proceed for 30 min at room temperature with 2 μg of poly(dI-dC) poly(dI-dC) (Amersham Bioscience), 20mM Tris-HCl (pH 7.9), 2mM $MgCl_2$, 50mM NaCl, 1mM EDTA, 10% glycerol, 0.1% Nonidet P-40, 1mM DTT, 50 μg/ml bovine serum albumin, the appropriate ^{32}P-labelled double-stranded (annealed) oligomer and, in some samples. After incubation, each sample was electrophoresed in a native 5% polyacrylamide gel using 0.25 x TBE buffer. The gels were dried and analyzed by using Bio-Imaging Analyzer (BAS 1000, Fuji Photo Films, Japan).

Analysis of c-Myc transcriptional regulatory activity in transient-cotransfection assay

For Analysis of c-Myc transcriptional regulatory activity, c-Myc expression vectors (pcDNA3-*CAM*1 or pcDNA3-*CAM*2) as effecter vector were cotransfected into KF cells with reporter vector (pGL3-MMBS-SV40) and pRL-SV40 using TransIT-LT1 (PanVera). After transfection, cells were placed in MEM-HEPES supplemented with 0.1% serum to reduce the activities of endogenous c-Myc for 48 hr. Cell extract preparations and dual luciferase assays were performed following the manufacturer's protocols (Toyo Ink).

Detectoin of mRNAs of c-Myc target genes

For c-Myc overexpression assay, the KF cells were transiently transfected at 20-30% confluence in 6-well dishes with 1 μg pcDNA3-*CAM*1 or pcDNA3-*CAM*2 using TransIT-LT1 (PanVera). After transfection, cells were placed in MEM-HEPES supplemented with 0.1% serum, to reduce the activities of endogenous c-Myc, for 24 hr before harvesting.

To inhibit expression of endogenous c-*myc*, the 25-mer morpholino antisense oligonucleotide (MO-*CAM*2) was purchased from Gene Tools, LLC (Philomath, Ore.). MO-*CAM*2 (5'-ACG CCA AAC TCG AAC T<u>CA T</u>CG GCA T-3') was designed against the 5'-untranslated region and starting codon (underlined sequences are complementary to starting codon) of *CAM*2. Transfections were performed by double-scrape delivery method following the manufacturer's protocols.

After harvesting, total RNA was isolated by using Sepasol RNA I Super (Nacalai Tesque), according to the manufacture's protocol. Expression of c-Myc target genes, *TERT*, *p53*, *Hsp70*, *ODC* and *cdc25* (for review, see ref. 6) were assessed by semiquantitative RT-PCR as described above. All cDNA fragments were newly isolated by degenerate PCR. Primer sequences were shown in Table I.

RESULTS

*Time course of induction of two c-*myc *mRNAs expression in carp cells--*

c-*myc* is the early-response gene. The transcription of c-*myc* is induced in several different cells by serum stimulation and causes the cell to exit G_0 and to proliferate (18, 19). Here, KF and KG cells were cultured for various time periods after serum stimulation. mRNA expression of *CAM*2 were induced in KF cells (P<0.001) and continued to increase during the first 12 h (Fig. 1A), but constant in KG cells (Fig. 1B). In contrast, the expression of *CAM*1 were significantly increased in KG cells (P<0.001) (Fig. 1B), but slightly changed in KF cells (P<0.05) (Fig. 1B). In mammalian cells, Myc protein itself inhibits *myc* transcription, and this negative feedback is thought to explain why the level of Myc declines from its initial peak to a lower steady value (20, 21, 22). However, carp c-*myc* mRNAs were not negatively autoregulated, which is different from mammalian c-*myc*.

*Promoter activity of the 5' flanking regions of carp c-*myc *genes--*

In order to analyze the structure of the carp c-*myc* promoters, we cloned 5' flanking regions of carp c-*myc* genes by inverse PCR. As a result, the nucleotide sequences of these clones (*CAM*1, 1578 bp and *CAM*2, 1235 bp) were determined (DDBJ, accession number AB103397 and AB103398). Using LALIGN algorithm (http://www.ch.embnet.org/software/LALIGN_form.html) (23), the nucleotide identities of these clones were low (45.3%). Using the BLAST program (24), there are no nucleotide identities between the 5' flanking regions of carp c-*myc* and other vertebrates (data not shown).

Various deletion mutants of these 5' upstream regions containing exon1 and intron1 of both c-*myc* were made and assayed for their promoter activity by dual lusiferase assay. The result is shown in Fig. 2. When the 5' deletion mutants were assayed, the exon1 of CAM1 had activities as a positive regulatory element. In *CAM*2, the regions of -891 to -1588 contained a negative regulatory element. Potential regulatory elements were analyzed using transcription factor binding site databases, TFSEARCH (http://www.cbrc.jp/research/db/TFSEARCH.html) (data not shown). The exon1 of *CAM*1, which showed higher promoter activity than that of *CAM*2, contained cdxA binding sites. cdxA protein can further activate transcription in cells in culture (25). In the upstream region between the -892 and –1588 of CAM2, the TFSEARCH program predicted potential binding sites for MZF1. MZF1 protein negatively regulates *CD3* and c-*myb* promoter activity in hematopoietic and non-hematopoietic cells upon binding to the MZF1 binding sites present in the 5'-franking region of both genes (26).

We next analyzed CpG island methylation of both c-*myc* in KF and KG cells by Southern blotting. *Bsa*HI digested 5' flanking regions of *CAM*1 in KG cells, although did not digest 5' flanking regions of both c-*myc* in KF cells (Fig. 3). Therefore, the promoter of *CAM*1 is active in KG at least. This result may correlate with induction of *CAM*1 mRNA expression in KG cells after serum stimulation (Fig. 1B).

Interaction of c-Myc with Max--

To test the specificity of the interaction between c-Myc and Max, we used GST fusion proteins of both Max and both c-Myc translated *in vitro*. CAM2 bound to both

Max but not to GST itself. In CAM1, a minimal level of binding to both Max was observed in this assay (Fig. 4A).

Two-hybrid system with carp cell lines confirmed the interaction of c-Myc and Max in another system. Both c-Myc are able to form a heterodimer with Max, but CAM2 binds to full-length Max more tightly than does CAM1 (Fig. 4B).

Binding of a Myc/Max complex to E-box--

Myc/Max complexes recognize the target sequence (E-box, CACGTG) and function as transcription activators. We assessed whether *in vitro* translated Myc could bind an E-box in a gel sift assay. Both c-Myc, along with its DNA binding partner Max, efficiently interacted with E-box. In competition assay, these Myc-DNA interactions were specifically inhibited by the addition of 50-fold excess amount of unlabeled competitor oligonucleotides containing the wild type (wt; CACGTG), but not the mutant (mt; CACGga), Myc binding site. Therefore, both c-Myc can bind to E-box with same affinity and specificity (Fig. 5).

Analysis of c-Myc transcriptional regulatory activity in transient-cotransfection assay in carp cells--

To investigate whether c-Myc is capable of transcriptional regulation of gene expression in live cells, we used a transient-transfection assay in which c-Myc expression vectors (pcDNA3-*CAM*1 or pcDNA3-*CAM*2) were cotransfected with a reporter plasmid (pGL3-MMBS-SV40). A single copy of the E-box is linked to an upstream of a SV40 promoter sequence and a luciferase gene in the plasmid. Transfection of c-Myc expression vectors led to increase in reporter gene activity.

Comparing CAM1 with CAM2, luciferase activity of CAM2 is partially but significantly higher than that of CAM1 (Fig. 6). High levels of luciferase activity were detectable when the reporter plasmid was transfected with control expression plasmid. Endogenous c-Myc may be contained in carp cells, because transcription of c-*myc* was not 0 % in reduced serum medium (refer hour 0 in Fig. 1A)

Transcriptional activation of c-Myc target genes--

To test the effect of each c-Myc overexpression, we analyzed the expression of target genes of c-Myc in KF cells transfected with expression vectors containing *CAM*1 or *CAM*2 cDNA. When pcDNA3-*CAM*1 was transiently tranfected, Hsp70, ODC, cdc25 mRNA expression was induced. In contrast, transfection of pcDNA3-*CAM*2 induced an increase in the expression of telomerase reverse transcriptase, *TERT*, in addition to *Hsp70*, *ODC* and *cdc25* (Fig. 7).

When the morpholino antisense oligonucleotide (MO-*CAM*2) was transfected, the expression level of *TERT* was suppressed (Fig. 7). Therefore, expression of *TERT* mRNA may be regulated by only CAM2. These results indicate that the two c-Myc proteins may have distinct abilities to induce activation of telomerase in carp cells. Other target genes were not suppressed by MO-*CAM*2, so CAM1 may be able to make up for the functions of CAM2. In this research, we could not design the morpholino antisense oligonucleotide against *CAM*1.

DISCUSSION

In the present study, we analyzed the biochemical functions of two distinct c-Myc in carp, and demonstrated several differences in the two c-Myc.

Here, we analyzed the expression of two c-*myc* genes in cultured cells of carp (KF and KG cells). *CAM*1 and *CAM*2 showed different expression after serum stimulation. The difference of expression observed in two cell lines may be concerned with the distinct properties between fibroblast (KF) and epithelial cells (KG). Next, to study the function of *cis*-acting elements of both c-*myc* genes, we carried out luciferase assay and analysis of the methylation of CpG islands. As a result, *CAM*1 and *CAM*2 showed differential transcription regulation systems. This result may reflect the difference of expression and the structure of exon1s and 5' franking regions (11).

Myc/Max heterodimer is able to function as transcription activators (27, 28). Although the expression pattern of *CAM*2 is similar to that of *Max*, *CAM*1 is not similar to it (12). The coordinate expression of *CAM*2 in the same tissues suggests that the CAM2/Max complex may serve an active physiological role as an original Myc/Max heterodimer throughout vertebrate evolution. Indeed, also in zebrafish, Myc/Max heterodimer as a regulator of gene expression has been maintained, and the expression of c-*myc* has been concomitant with *Max* expression (14). Affinity *in vitro* and *in vivo* between Max and two c-Myc showed that CAM2 binds to Max more tightly than does CAM1. This result may correlate with our previous report (12). The difference of affinity between Max and two c-Myc in carp cells may influence transcriptional activity to target genes.

Expressions of c-Myc target genes following overexpression of each c-Myc and the knock-down of *CAM*2 showed that Hsp70, ODC and cdc25 mRNA expression were

regulated by both c-Myc. Gel-sift assay revealed both c-Myc could bind to the same target sequences, E-box, indicating that CAM1 and CAM2 have the same target gene. Therefore, Overlapping functions between both c-Myc raise the possibility of complementation of CAM2 deficiency by CAM1. But in *TERT*, that is a subunit of telomerase, the transcription of this gene was regulated by only CAM2. This result may be related to the fact that transcriptional activity of CAM2 is higher in some parts than that of CAM1.

Our present data demonstrate that *CAM*1 and *CAM*2 do not have completely redundant functions, and have distinct properties in terms of expression patterns, transcription regulation systems, cooperation with Max, and transcriptional activation. In other words, it is likely that *CAM*1 and *CAM*2 have overlapping but somewhat distinct functions. Moreover, it is noteworthy that CAM1, but not CAM2, has apparently deviated from the original function of c-*myc* gene through vertebrate evolution. This result agrees with the hypothesis that one of the 2 duplicated genes may evolve faster to obtain a new function or become silent, being subsequent to the tetraploidization event (1). According to phylogenetic analysis, *CAM*1 is evolving 1.6 times faster than *CAM*2 after gene duplication, and *CAM*2 is conserved throughout vertebrate evolution (10). We reasoned from the results of these phylogenetic and functional analyses that *CAM*2 was an ortholog of mammalian c-*myc* and *CAM*1 was a novel homolog. Our results showed that the *CAM*1, which is evolving faster, might obtain a new function different from c-*myc* and regulate transcription of the target gene by changing the expression level in tissues and cells, supposing that *myc* family has been formed in fish as well as mammalian.

Although we still do not know the molecular mechanisms that define the distinct ability between both c-Myc, it is likely that some factors interact specifically with each of the two c-Myc proteins, eg. other proteins except tested ones that may effect on the transcriptional activity of each of the two c-Myc. Studying macromolecular interactions in a signaling pathway is the key to know the biological function. Indeed, recent advances have shown that Myc collaborates with a variety of other cellular factors at both of N- and C-terminal domains to mediate its many biological activities (for review, see ref. 6). We are currently investigating the mechanisms that underlie the difference between CAM1 and CAM2 in their cooperatively with new Myc-interacting proteins.

ACKNOWLEDGMENTS

We thank Dr. Ronald P. Hedrick (University of California, Davis) for gifts of KF and KG cells.

REFERENCES

1. Ohno, S. (1970) Evolution by gene duplication. Springer Verlag Press, Heidelberg, New York.
2. Lundin, L. G. (1993) Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. *Genomics* 16, 1-19
3. Allendorf, F. W. and Thorgaard, G. H. (1984) Tetraploidy and the evolution of salmonid fishes, in "Evolutionary Genetics of Fishes" (ed. By Turner, B. J.). Plenum Pres, New York and London, pp. 1-53
4. Ohno, S. (1993) Patterns in genome evolution. *Curr. Opin. Genet. Dev.* 3, 911-914
5. Roy, A. L., Carruthers, C., Gutjahr, T. and Roeder, R. G. (1993) Direct role for Myc in transcription initiation mediated by interactions with TFII-I. *Nature* 365, 359-361
6. Oster, S. K., Ho, C. S. W., Soucie, E. L. and Penn, L. Z. (2002) The *myc* oncogene: MarvelouslY Complex. *Adv. Cancer Res* 84, 81-154
7. Luscher, B. (2001) Function and regulation of the transcription factors of the Myc/Max/Mad network. *Gene*, 277, 1-14
8. Ryan, K. M. and Birnie, G. D. (1996) *Myc* oncogenes: the enigmatic family. *Biochem. J.* 314, 713-721

9. Zhang, H., Okamoto, N. and Ikeda, Y. (1995) Two c-*myc* genes from a tetraploid fish, the common carp (*Cyprinus carpio*). *Gene* 153, 231-236

10. Zhang, H. (1994) Isolation and DNA Sequence of one c-*myc* Gene from Goldfish *Carassius auratus* and the Molecular Phylogeny for c-Myc of Vertebrates. In: Doctoral dissertation, Tokyo University of Fisheries.

11. Futami, K., Komiya, T., Zhang, H. and Okamoto, N. (2000) Determination of heterogeneous transcription start points of two c-*myc* genes from the common carp (*Cyprinus carpio*). *Gene* 245, 43-47

12. Futami, K., Komiya, T., Zhang, H. and Okamoto, N. (2001) Differential expression of max and two types of c-*myc* genes in a tetraploid fish, the common carp (*Cyprinus carpio*). *Gene* 269,113-119

13. Blackwood, E. M. and Eisenman R. N. (1991) .: Max: A Helix-Loop-Helix Zipper Protein That Forms a Sequence-Specific DNA-Binding Complex with Myc. *Science* 251, 1211-1217

14. Schreiber-Agus, N., Horner, J., Torres, R., Chiu, F. C. and Depinho, R. A. (1993) Zebra Fish *myc* and *max* Genes: Differential Expression and Oncogenic Activity throughout Vertebrate Evolution. *Mol. Cell. Biol.* 13, 2765-2775

15. Fijan, N., Sulimanoviæ D., Bearzotti, M., Muziniæ D., Zwillenberg, L. O., Chilmonczyk, S., Vautherot, J. F. and de Kinkelin, P. (1983) Some properties of the *Epithelioma papulosum cyprini* (*EPC*) cell line from carp *Cyprinus carpio*. *Ann. Virol.* 134 E, 207-220

16. Katagiri, T., Hirono, I. and Aoki, T. (1997) Identification of a cDNA for medaka

17. cytoskeletal β-actin and construction for the reverse transcriptase- polymerase chain reaction (RT-PCR) primers. *Fisheries Science* 63, 73-76

18. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl. K. (1992) Current Protocols in Molecular Biology, John Wiley & Sons, New York

19. Kelly, K., Cochran, B. H., Stiles C. D. and Leder P. (1983) Cell-specific regulation of the c-*myc* gene by lymphocyte mitogens and platelet-derived growth factor. *Cell* 35, 603-610

20. Greenberg, M. E. and Ziff, E. B. (1984) Stimulation of 3T3 cells induces transcription of the c-*fos* proto-oncogene. *Nature* 311, 433-438

21. Penn, L. J., Brooks, M. W., Laufer, E. M. and Land, H. (1990) Negative autoregulation of c-*myc* transcription. *EMBO J.* 9, 1113-1121.

22. Facchini, L. M., Chen, S., Marhin, W. W., Lear, J. N. and Penn, L. Z. (1997) The Myc negative autoregulation mechanism requires Myc-Max association and involves the c-*myc* P2 minimal promoter. *Mol. Cell. Biol.* 17, 100-114.

23. Oster, S. K., Marhin, W. W., Asker, C., Facchini, L. M., Dion, P. A., Funa, K., Post, M., Sedivy, J. M. and Penn, L. Z. (2000) Myc is an essential negative regulator of platelet-derived growth factor beta receptor expression. *Mol Cell Biol.* 20, 6768-78

24. Huang, X. and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.* 12, 337-357.

25. Altschul, S. F., Gish, W., Miller, W., Myers, E. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410

26. Margalit, Y., Yarus, S., Shapira, E., Gruenbaum, Y. and Fainsod, A. (1993) Isolation and characterization of target sequences of the chicken CdxA homeobox gene. *Nucleic Acids Res.* 21, 4915-4922.

27. Perrotti, D., Melotti, P., Skorski, T., Casella, I., Peschle, C. and Calabretta, B. (1995) Overexpression of the zinc finger protein MZF1 inhibits hematopoietic development from embryonic stem cells: correlation with negative regulation of CD34 and c-*myb* promoter activity. *Mol. Cell. Biol.* 15, 6075-6087.

28. Amati, B., Dalton, S., Brooks, M. W., Littlewood, T. D., Evan, G. I. and Land, H. (1992) Transcriptional activation by the human c-Myc oncoprotein in yeast requires interaction with Max. *Nature* 359, 423-429

29. Kato, G. J., Lee, W. M., Chen, L. L. and Dang, C. V. (1992) Max: functional domains and interaction with c-Myc. *Genes. Dev.* 6, 81-92.

Figure legend

Fig. 1. Time course of induction of two c-*myc* mRNAs expression in carp cells. *A*, Serum induced c-*myc* expression in KF cells. *B*, Serum induced c-*myc* expression in KG cells. The dotted line with the open circles shows the *CAM*1 mRNA expression, whereas the closed squares show the *CAM*2 mRNA expression. Results shown are the mean ± SE (n = 3). Significant levels at $P<0.05$ (*) and $P<0.001$ (***) with ANOVA are indicated, whereas NS denote the levels not significant at $P>0.05$.

Fig. 2. Promoter activity of the 5' flanking regions of carp c-*myc* genes. Relative luciferase activities in KF cells are shown by open bars, while in KG cells, are shown by closed bars. Results shown are the mean ± SD (n = 3).

Fig. 3. Analysis of the methylation of CpG islands of c-*myc* genes. Arrows show the bands detected in only KG cells.

Fig. 4. Interaction of c-Myc with Max. *A*, *In vitro* binding of c-Myc with Max. Products from the *in vitro* translation reaction are shown as input. Max1 represent the full-length type. Max2 represent the spliced variant. *B*, Analysis of c-Myc binding to Max by a two-hybrid assay using fish culture cells. Results shown are

the mean ± SD (n = 3).

Fig. 5. Binding of a Myc/Max complex to E-box. Unlabeled CAM1, CAM2 and Max proteins were translated *in vitro*. Each reaction contained 1 ng of ^{32}P-labeled probe. The unlabeled competitor probe (wt or mt) was added in 50-fold excess (50 ng). Asterisks (*) indicates a specific DNA-protein complex. RL, reticulocyte lysate; wt, wild type; mt, mutant.

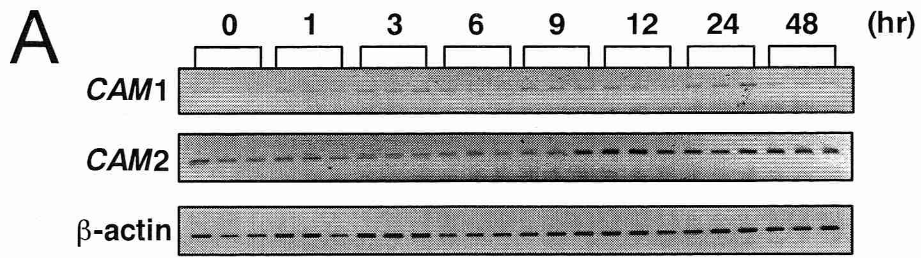
Fig. 6. Analysis of c-Myc transcriptional regulatory activity in transient-cotransfection assay. Results shown are the mean ± SD (n = 3).

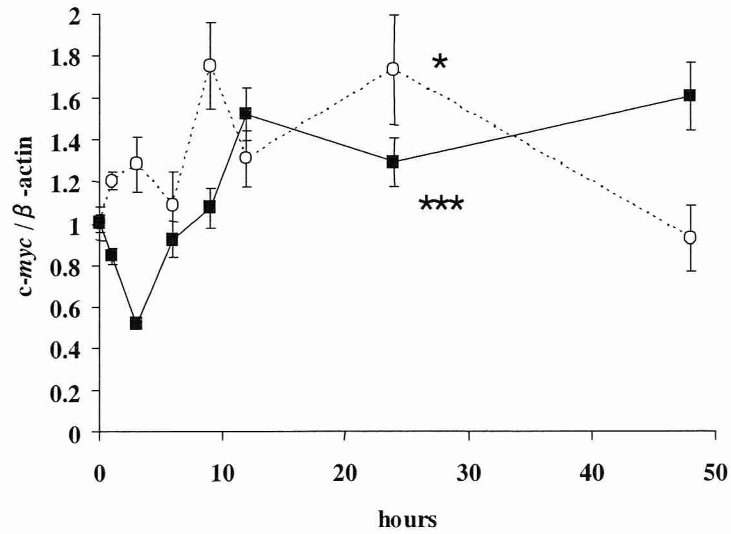
Fig. 7. Detections of mRNAs of c-Myc target genes. *A*, The effect of each c-Myc overexpression. *B*, The effect of morpholino antisense *CAM2*. Expression of cytoskeletal β-actin gene was used for internal control. Experiments were performed in duplicate. Table I. Primers used for RT-PCR.

Table I. Primers used for RT-PCR.

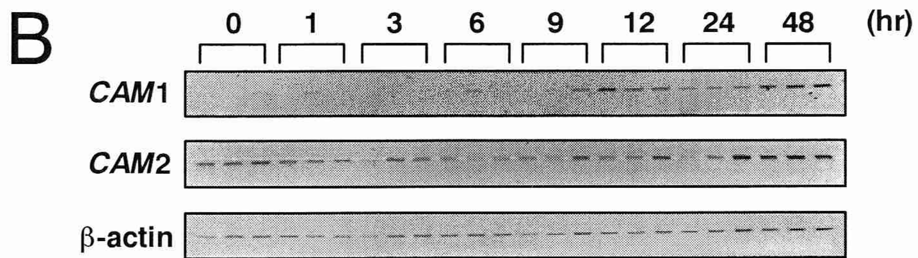
Gene	Primer sequences	Annealing templature (℃)	Cycles
TERT	TGA AGG GAT CGC TAA AGG AGG	57	35
	GGT CCA GAA AGA TGT CGG TAC		
p53	GTG TCT GTG GAT ATA CTG GTG G	57	25
	CCT ATT CAT CCC ACC CAT ACA G		
HSP70	ACG TCC TGA TCT TTG ACC TG	57	30
	GTC CAT CTT GGG TCT CTC AG		
ODC	TAY GCI AAY CCI TGY AAR CAR G	51	30
	ACI GTR TAI GCI CCC ATR TTY TC		
cdc25	GAY TGY MGI TAY CCI TAY GAR TA	51	30
	RAA RAA YTC YTT RTA ICC ICC		

A





B



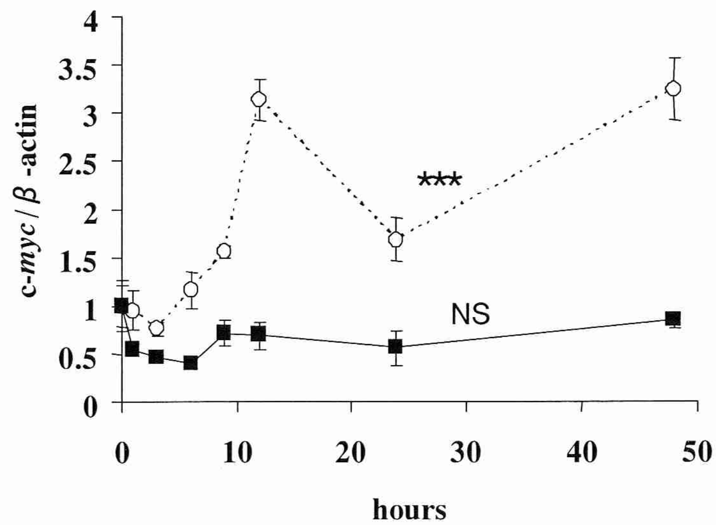


Fig. 1 (Futami)

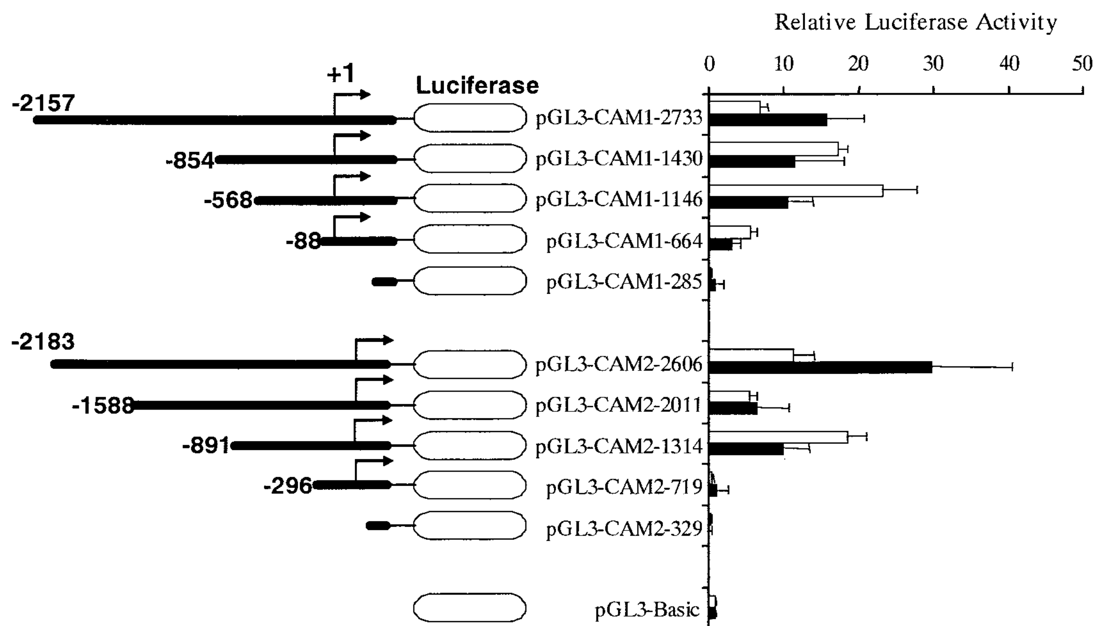


Fig. 2 (Futami)

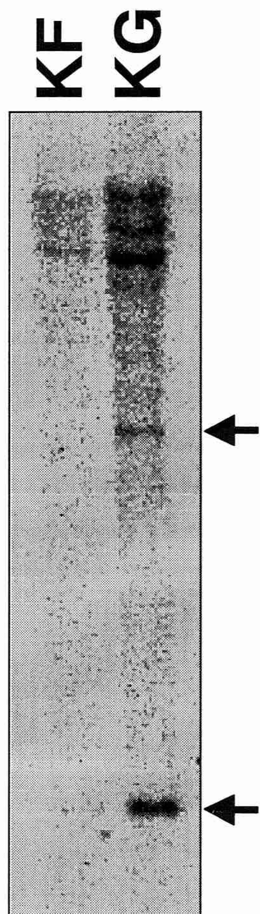
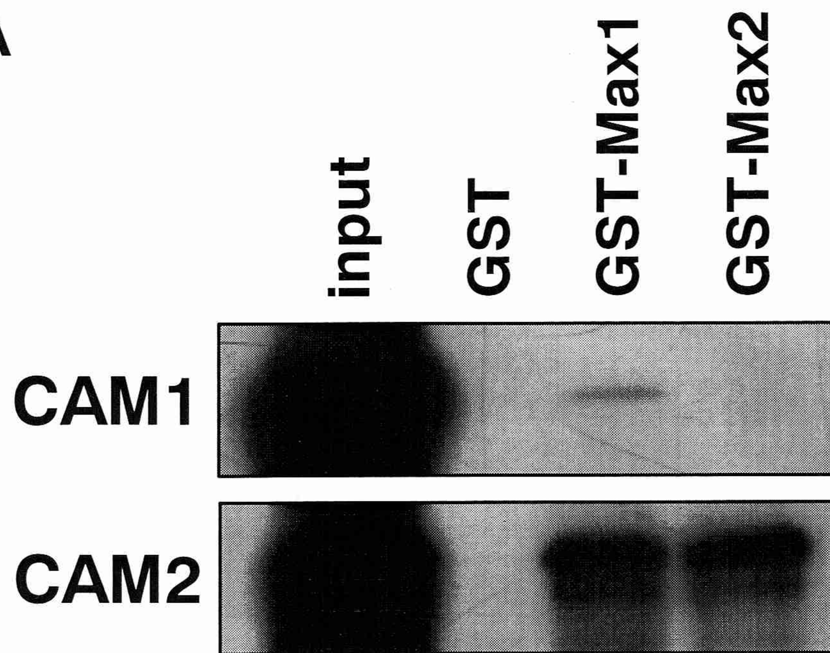


Fig. 3 (Futami)



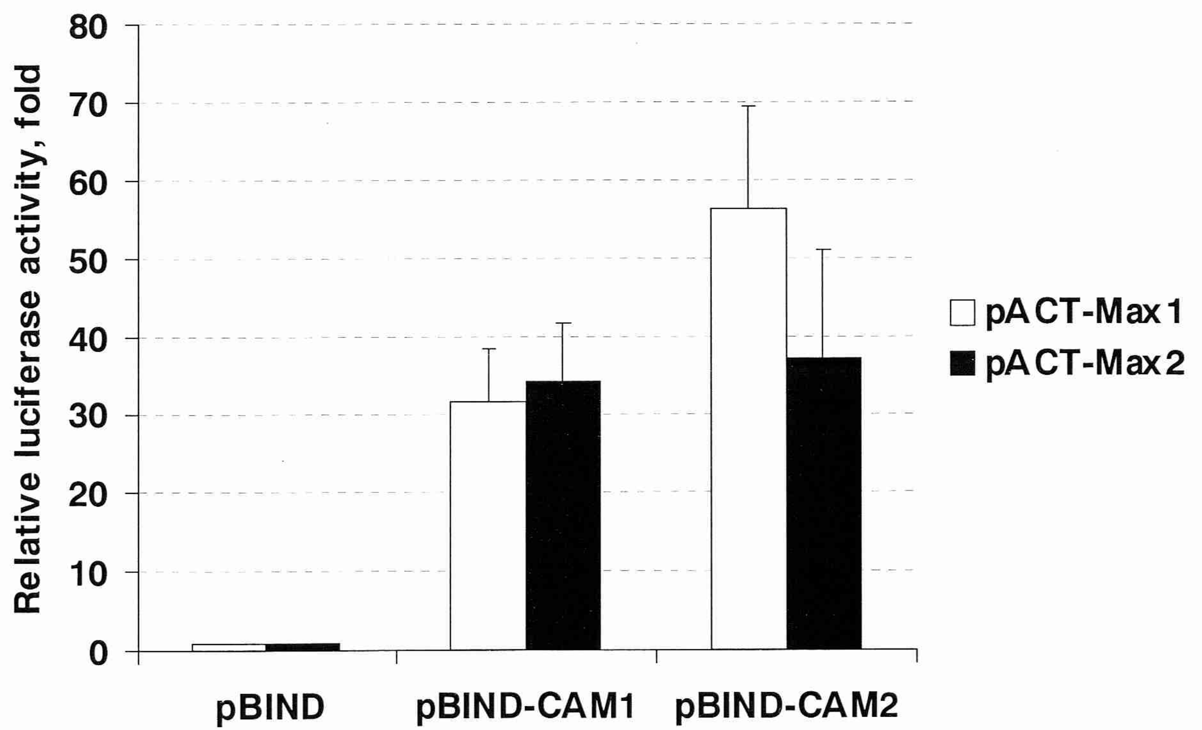


Fig. 4 (Futami)

Protein

RL	+	+	+	+	+	+	+	+	+	+
CAM1			+	+	+	+				
CAM2							+	+	+	+
Max		+		+	+	+		+	+	+

Probe	+	+	+	+	+	+	+	+	+	+
wt Competitor					+				+	
mt Competitor						+				+

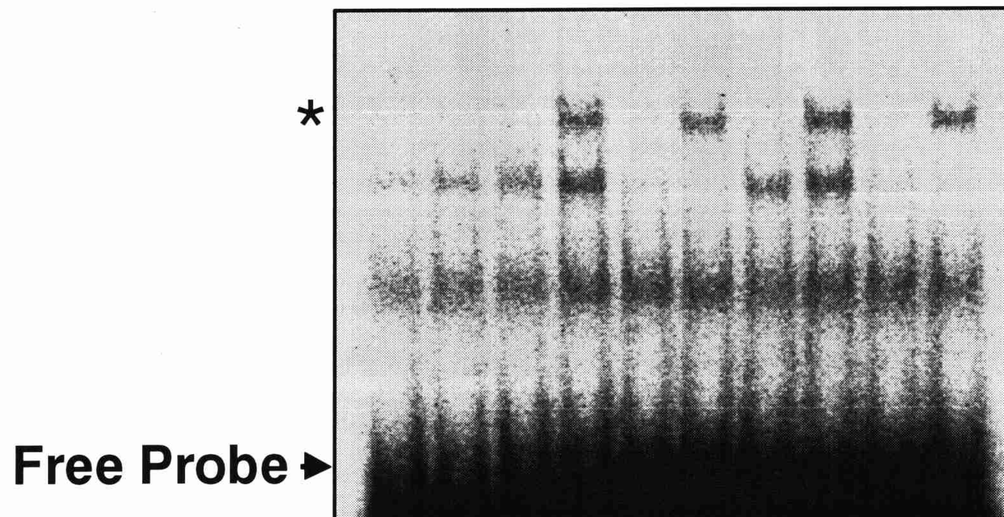


Fig. 5 (Futami)

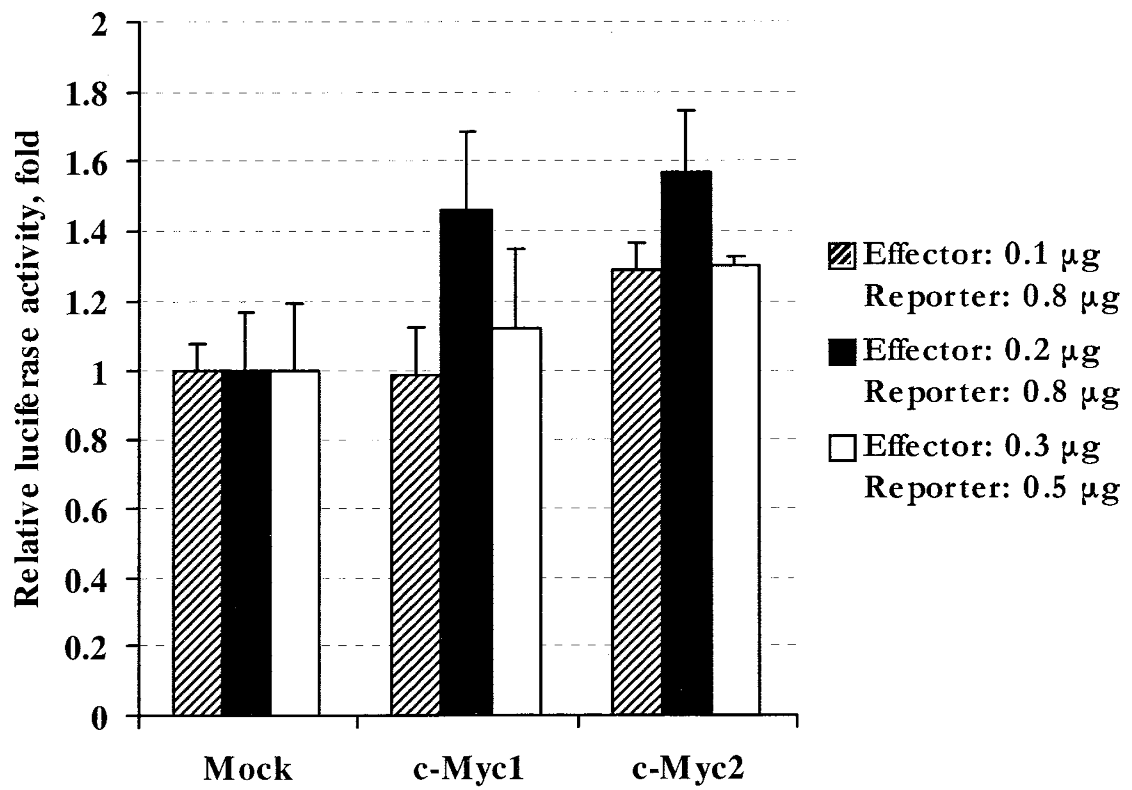


Fig. 6 (Futami)

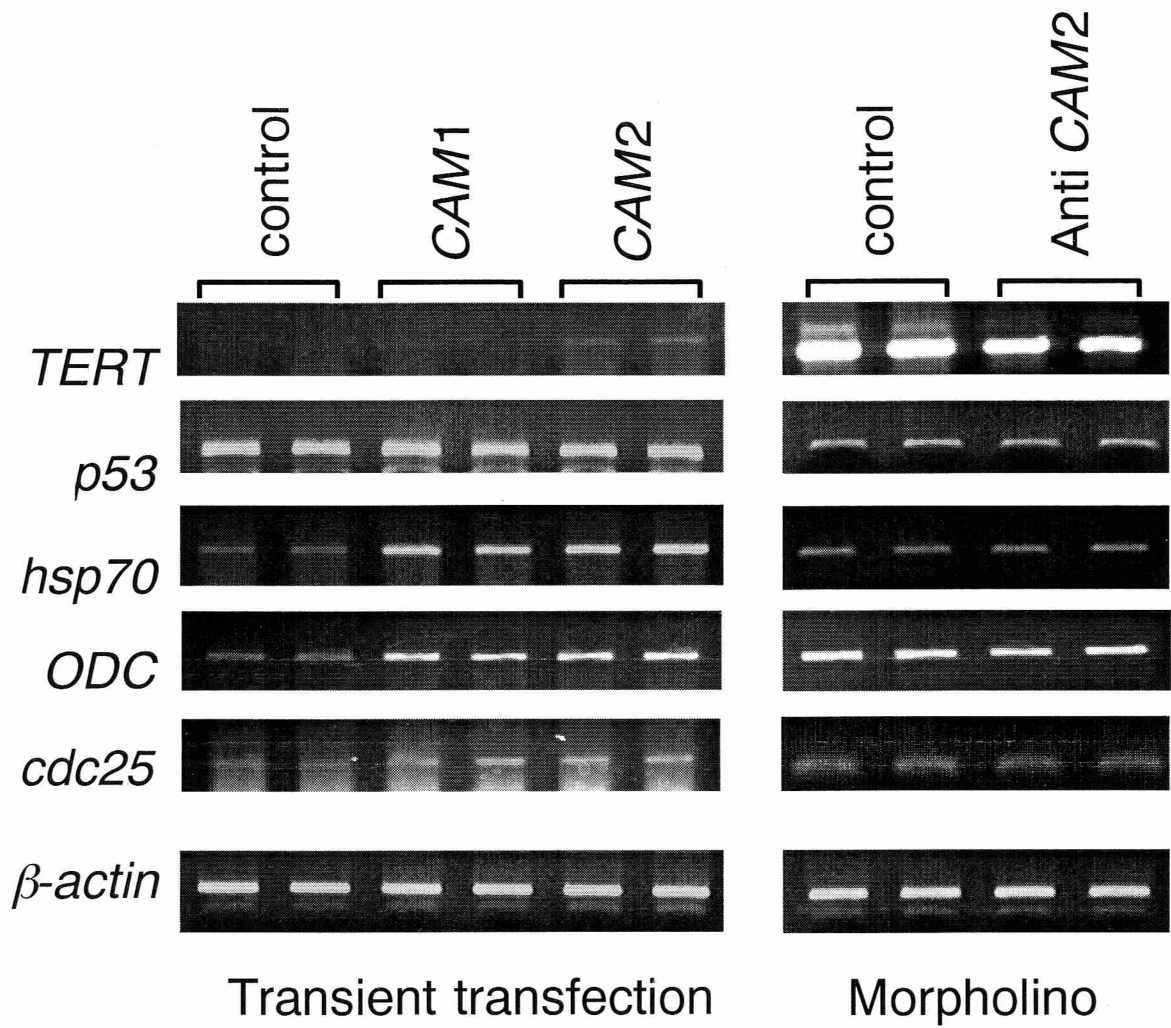


Fig. 7 (Futami)

第5章　Bacterial Two-Hybrid Systemによる新規c-Myc結合蛋白質の網羅的スクリーニング

Genome-Wide Screening of New c-Myc Interacting Proteins using Bacterial Two-Hybrid System

Keywords: cDNA library; protein-protein interaction; ribosomal proteins; tetraploid

Kunihiko Futami [a], Huan Zhang [b], Nobuaki Okamoto [a,*]

[a] *Department of Aquatic Biosciences, Tokyo University of Fisheries, Konan 4, Minato-ku, Tokyo 108-8477, Japan*

[b] *Department of Marine Sciences, University of Connecticut, Groton, CT06340, USA*

*Corresponding author. Tel: +81 3 5463 0547, Fax: +81 3 5463 0552; e-mail: nokamoto@tokyo-u-fish.ac.jp

Abbreviations: *CAM*1, carp c-*myc*1 encoding c-Myc1; *CAM*2, carp c-*myc*2 encoding c-Myc2; cDNA, DNA complementary to RNA; c-*myc*, gene(s) encoding c-Myc; c-Myc, cellular Myc; nt, nucleotide(s).

Abstract

The proto-oncogene c-*myc* is thought to be one of the most important genes in controlling cell proliferation. However, the mode of action and its interaction with the signaling pathway is still unclear. In this research, we identified new c-Myc interacting proteins using bacterial two-hybrid system. As a result, 10 clones were cited as candidates of CAM1 specific interacting proteins and 11 clones were cited as candidates of CAM2 specific interacting proteins. The difference of interacting proteins between these two c-Myc suggests that CAM1 and CAM2 may evolve to acquire different functions in signaling pathways after the tetraploid event. The present study may help us to understand the biological roles of the c-*myc* genes in carp, besides knowing another function of c-*myc*, which is unknown in mammalian.

1. Introduction

Studying macromolecular interactions in a signaling pathway is the key to know the biological function. Indeed, the identification of Max as a required partner of Myc proteins was pivotal and at the time led to a revolution in understanding of Myc function. Until recently, no other protein interactors that are so intimately linked to Myc activity had been identified. Recent advances have shown that Myc collaborates with a variety of other cellular factors at both its N- and C-terminal domain to mediate its many biological activities (Oster *et al.*, 2002). In this research, we identified new c-Myc interacting proteins using bacterial two-hybrid system.

2. Materials and methods

2.1. Construction of carp cDNA library for "target" plasmid

Total RNA was isolated from KF cells established from the common carp, *Cyprinus carpio*, by using Sepasol RNA I Super (Nacalai Tesque), according to the manufacture's protocol. Poly(A)+ mRNA was purified by Oligotex-dT30<Super> mRNA Purification Kit (TAKARA BIO INC., Otsu, Japan) and reverse transcribed

using the TimeSaver cDNA synthesis kit (Amersham Bioscience). The obtaind cDNA with *Eco*RI-*Xho*I adapter was cloned into the identical restriction site of pTRG target vector (Stratagene) to construct a library.

2.2. Bacterial two-hybrid screening

The BacterioMatch Two-Hybrid System Vector Kit was purchased from Stratagene. The cDNAs encoding both c-*myc* were amplified by PCR-based strategy. The PCR products, which were produced by an additional *Eco*RI site in 5'-upstream and *Bam*HI site in 3'-downstream, were subcloned in-frame into the identical restriction sites of pBT to generate "bait" plasmids, pBT-*CAM*1 and pBT-*CAM*2. Both bait and target plasmids were introduced into BacterioMatch two-hybrid system reporter strain competent cells (Stratagene), using manufacture's protocol. Transformed cells were plated on LB-CTCK agar plates containing 250 μg/ml carbenicillin, 15 μg/ml tetracycline, 34 μg/ml chloramphenicol and 50 μg/ml kanamycin, and incubated at 30°C for 24 h. Carbenicillin resistant colonies were picked up and assayed for *lacZ* activity using a β-Galactosidase reporter to validate specificity of protein-protein interactions. The nt sequences of the positive cDNA clones were analyzed by dye terminator cycle sequencing using the ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Cetus, USA)

and nucleotide sequence databases were searched for homologous sequences using the BLASTN and BLASTX programs.

3. Results and discussion

Using CAM1 and CAM2 as bait plasmids, we screened a cDNA library of 2 x 10^5 colonies three times respectively. Using CAM1 as bait, 16 positive colonies were obtained from cDNA library, while using CAM2, 15 positive colonies were obtained. Sequencing of the inserted cDNAs revealed that all clones had not been reported as Myc-interacting proteins in other vertebrates. As a result, 10 clones were cited as candidates of CAM1 specific interacting proteins (Table 1) and 11 clones were cited as candidates of CAM2 specific interacting proteins (Table 2). The difference of interacting proteins between these two c-Myc suggests that CAM1 and CAM2 may evolve to acquire different functions in signaling pathways after the tetraploid event. Furthermore, the CAM1, which is evolving faster, may obtain new functions different from c-Myc. However, the biological significant of these proteins in carp cells is unclear at present. Interestingly, some ribosomal proteins were identified as both c-Myc interacting proteins. Intracellular localization of ribosomal proteins is different

from that of c-Myc proteins, so the interactions between these proteins are unlikely. However, recent reports on the involvement of ribosomal proteins in various genetic diseases and studies on the "extraribosomal functions" of these proteins have cast some light on their localization and functions (Wool et al., 1996). Ribosomal proteins are associated with cell differentiation and malignant tumorigenesis, and regulation of ribosome-interacting proteins seems to be tightly associated with the stress response, apoptosis and carcinogenesis. c-Myc is also associated with all of these phenotypes. Furthermore, c-Myc induces the transcription and translation of 40S and 60S ribosomal proteins (Coller et al., 2000, Guo et al., 2000, Boon et al., 2001, Neiman et al., 2001, Schuhmacher et al., 2001, Shiio et al., 2002). Therefore, interaction between c-Myc and ribosomal proteins may be considered as a possibility. The present study may help us to understand the biological roles of the c-*myc* genes in carp, besides knowing another function of c-*myc*, which is unknown in mammalian. However, false positives are often detected in two-hybrid system in general, so further studies are needed to determine the protein-protein interaction between two c-Myc and these proteins *in vivo*.

Acknowledgement

We thank Dr. Ronald P. Hedrick (University of California, Davis) for gifts of KF cells. This work was supported in part by Research Fellowships of the Japan Society for the Promotion of Science (JSPS) for Young Scientists.

References

Boon, K., Caron, H. N., van Asperen, R., Valentijn, L., Hermus, M. C., van Sluis, P., Roobeek, I., Weis, I., Voute, P. A., Schwab, M. and Versteeg, R., 2001. N-myc enhances the expression of a large set of genes functioning in ribosome biogenesis and protein synthesis. EMBO J. 20, 1383-1393.

Coller, H. A., Grandori, C., Tamayo, P., Colbert, T., Lander, E. S., Eisenman, R. N. and Golub, T. R., 2000. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. Proc. Natl. Acad. Sci. USA 97, 3260-3265.

Guo, Q. M., Malek, R. L., Kim, S., Chiao, C., He, M., Ruffy, M., Sanka, K., Lee, N. H., Dang, C. V. and Liu, E. T., 2000. Identification of c-myc responsive genes using rat cDNA microarray. Cancer Res. 60, 5922-5928.

Neiman, P. E., Ruddell, A., Jasoni, C., Loring, G., Thomas, S. J., Brandvold, K. A., Lee, R., Burnside, J. and Delrow, J., 2001. Analysis of gene expression during myc oncogene-induced lymphomagenesis in the bursa of Fabricius. Proc. Natl. Acad. Sci. USA 98, 6378-6383.

Oster, S. K., Ho, C. S. W., Soucie, E. L. and Penn, L. Z., 2002. The *myc* oncogene: MarvelouslY Complex. Adv. Cancer Res 84, 81-154.

Schuhmacher, M., Kohlhuber, F., Holzel, M., Kaiser, C., 2001. Burtscher, H., Jarsch, M., Bornkamm, G. W., Laux, G., Polack, A., Weidle, U. H. and Eick, D., The transcriptional program of a human B cell line in response to Myc. Nucleic Acids Res. 29, 397-406.

Shiio, Y., Donohoe, S., Yi, E. C., Goodlett, D., Aebersold, R. and Eisenman, R. N., 2002. Quantitative proteomic analysis of Myc oncoprotein function. EMBO J. 21, 5088-5096.

Wool, I. G., 1996. Extraribosomal functions of ribosomal proteins. Trends. Biochem. Sci. 21, 164-165.

Table I. Candidates of CAM1 specific interacting proteins.

Genes	Accession no.	Species	E-value
unknown			
ribosomal protein L9	**AF401562**	**channel catfish**	**3e-31**
guanine monphosphate synthetase (GMPS)	**XM_010978**	**human**	**3e-12**
ribosomal protein L17	AF401571	channel catfish	6e-60
40S ribosomal protein S10	AF402818	channel catfish	e-113
40S ribosomal protein S10	AF402818	channel catfish	4e-67
40S ribosomal protein S16	**AF402825**	**channel catfish**	**3e-37**
ribosomal protein L41 mRNA	AF503957	orange-spotted grouper	2e-28
ribosomal protein L13a	**AF401568**	**channel catfish**	**2e-93**
40S ribosomal protein S30	**AAK95215**	**channel catfish**	**1e-25**
40S ribosomal protein S5 (rps5)	**AF506223**	**zebrafish**	**0.0**
MRPb and MASPb genes for mannose-binding lectin-associated serine protease (MASP) and MASP-related protein	**AB030447**	**carp**	**0.002**
40S ribosomal protein S29 mRNA	AF402840	channel catfish	1e-40
ribosomal protein L3 mRNA	AF401554	channel catfish	1e-36
ribosomal protein S12 mRNA	**U94500**	**Tilapia nilotica**	**e-141**
ribosomal protein S12 mRNA	**U94500**	**Tilapia nilotica**	**e-141**

Bold letters show the proteins interacted with only CAM1.

Table II. Candidates of CAM2 specific interacting proteins.

Genes	Accession no.	Species	E-value
ubiquitin-like fusion protein An1b	**JN0674**	**Xenopus laevis**	**4.8**
40S ribosomal protein S10	AF402818	channel catfish	e-169
unknown			
collagen, type IV, alpha 1 (COL4A1)	**XM_049913**	**human**	**2.7**
ubiquitin A-52 residue ribosomal protein fusion product 1	**AAK31162**	**human**	**9e-18**
Ictacalcin	**Q91061**	**channel catfish**	**1e-05**
cDNA clone 1-16	**AJ009316**	**Xenopus laevis**	**2e-29**
ribosomal protein L19	**AF401574**	**channel catfish**	**9e-28**
unknown			
complete mitochondrial genome	**X61010**	**carp**	**e-163**
Cytochrome c oxidase polypeptide VIIb, mitochondrial precursor (IHQ)	**P13183**	**cow**	**1e-09**
40S ribosomal protein S29	AF402840	channel catfish	4e-47
40S ribosomal protein S9	**AAK95191**	**channel catfish**	**4e-19**
60S RIBOSOMAL PROTEIN L3 (L4)	P21531	Norway rat	3e-26
ribosomal protein L41 mRNA	AF503957	orange-spotted grouper	4e-29

Bold letters show the proteins interacted with only CAM2.

総合考察

本研究は、Ohno (1970) が提唱した「倍数性進化」の仮説を、4倍性魚類であるコイの c-*myc* 遺伝子に着目して実証しようとするものである。

コイ c-*myc* 遺伝子の転写開始点をオリゴキャップ法 (Maruyama and Sugano, 1994) により解析し、魚類の c-*myc* 遺伝子に第1エキソンが存在することを初めて明らかにした。*CAM1* と *CAM2* の第1エキソンを比較したところ、その相同性は他のエキソンよりも低く、転写開始点の位置や数も異なっており、*CAM1* と *CAM2* の間には何らかの機能分化が起きている可能性が示された（第2章）。

Neighbor-Joining 法による分子系統樹から、*CAM1* と *CAM2* はそれぞれ進化速度が異なり、*CAM1* は *CAM2* よりも進化速度が 1.6 倍速いことが分かっている(張、1994)。しかし、*CAM1* と *CAM2* の機能や発現の違いについては、これまで検討されてこなかった。そこで RT-PCR により、*CAM1* と *CAM2* の発現量を組織、および培養細胞株で調べたところ、この 2 タイプの c-*myc* 遺伝子の発現パターンに違いが認められた。またゼブラフィッシュでは、ヒトと同様に c-*myc* 遺伝子が、Myc 蛋白質とヘテロダイマーをつくる Max の遺伝子の発現と共同歩調をとることが明らかにされている (Schreiber-Agus *et al.*, 1993)。そこで、*Max* の cDNA をクローニングし、ノーザンブロット解析を行ったところ、*Max* は *CAM2* と同様の発現パターンを示したが、*CAM1* とは異なった。これらのことから、*CAM1* は c-*myc* としての機能とは別の新しい機能を持つように進化している可能性が強く示唆された（第3章）。

そこで、2 タイプ の c-Myc が蛋白質レベルでどのような挙動を示すかについて明らかにするために、コイおよびニシキゴイ由来培養細胞株を用いて 2 タイプの c-Myc の生化学的機能の違いについて解析をおこなった。その結果、CAM1 と CAM2 の機能はオーバーラップしてはいたものの、Max に対する結合特異性は異なり、転写活性にも差が認められた（第 4 章)。さらに、*CAM1* が *CAM2* と異なる発現パターンをし、別の蛋白質と相互作用していることは、*CAM1* が新しい機能を獲得しつつあるということを示している（第 4 章および第 5 章)。一般に、重複によって生じた遺伝子のうち、あるものは新しい機能を獲得し、またあるものは機能を失い、偽遺伝子となる。コイにおいても例外ではなく、現在、コイの重複した遺伝子の半分が機能を失ったと推定されている (Ferris and Whitt, 1977)。しかし、*CAM1* はいまでも機能を失っていない。進化速度の速い *CAM1* は発現する組織や細胞を変え、種々の結合蛋白質との組み合わせで標的遺伝子を変化させているのではないだろうか。このことは、魚類においても、哺乳類にみられるような *myc* ファミリーが形成されつつあることを想定させ、Ohno の仮説を支持するものである。

本研究では、c-*myc* の生物学的役割を明らかにするために、それぞれの c-*myc* を一過性に導入した細胞と *CAM2* をノックダウンした細胞を用いて、TRAP (Telomeric Repeat Amplification Protocol) 法によるテロメラーゼ活性の測定、および MTT 法による細胞増殖活性の測定をおこなったが、いずれもコントロールに対して有意な差は見られなかった (data not shown)。生物学的機能といった場合、それは 1 個の蛋白質によって担われているわけではなく、複数の要素によって形成されるネットワークによって担われていると考えられ、したがって、2 タイプの

c-*myc*の生物学的な意義を考える際には、それぞれのc-Mycがどのネットワークの中ではたらいているか、また、そのネットワークの中のどこに位置するかという情報が必要になる。

今後の課題として、(1) Bacterial two-hybrid systemで得られた蛋白質が実際に*in vivo*でも相互作用を示すかどうかを免疫沈降による共沈により確認すること、(2) 2タイプのc-Mycのそれぞれの標的遺伝子をクローニングすることなどが残されている。本研究での結果は、倍数性化後の重複遺伝子の進化を機能との関連で研究する上で新しい知見を与えただけでなく、脊椎動物における*myc*ファミリーの形成の解明にも広く貢献するものと期待でき、癌遺伝子でもある*myc*の機能を進化の側面から明らかにすることにつながる可能性を含んでいる。本研究は、魚類のc-*myc*遺伝子の進化および機能に関する研究に、新たな展開をもたらすものであると確信する。

引用文献

Ferris, S. D., Whitt, G. S.: *Experientia*, **33**, 1299-1301 (1977)

Maruyama, K. and Sugano, S.: *Gene*, **138**, 171-174 (1994)

Ohno, S.: Evolution by Gene Duplication. Springer Verlag Press, Heidelberg, New York, 1970

Schreiber-Agus, N., Horner, J., Torres, R., Fung-Chow Chiu and Depinho, R. A.: *Mol. Cell. Biol.*, **13**, 2765-2775 (1993)

張寰：東京水産大学博士論文、1994

謝辞

本研究を行うに当たり、終始懇切なご指導を賜り、また本論文のご校閲を頂いた東京水産大学 資源育成学科 水族生理学研究室 岡本 信明 教授、舞田 正志 助教授、水族病理学研究室 福田 穎穂 教授ならびに遺伝生化学講座 青木 宙 教授に深甚たる謝意を表します。また本研究を行うに当たり、ご指導、ご助言頂いた University of Connecticut 張 寰 博士ならびに雪印乳業 小宮 猛 氏に心から感謝します。KF および KG 細胞は University of California, Davis, Dr. Ronald P. Hedrick から供与されました。東京水産大学 放射性同位元素利用施設、伊藤由加里さんならびに高野和輝さんには多くの便利を提供して頂いた。特に期して衷心より感謝の意を述べたい。

また、水族生理学研究室の皆様には終始迷惑のかけ通しであったことをお詫びするとともに、彼らおよび彼女らの支えがなければ、この研究を成し遂げることはできなかったことを最後に記し、心から厚く感謝申し上げます。

本研究の一部は、日本学術振興会の特別研究員奨励費によって行われたものである。

資料

総説

二見邦彦・張寰・岡本信明：4倍性魚類コイの c-*myc* 遺伝子2タイプの進化．

“小型魚類研究の新展開―脊椎動物の発生・遺伝・進化の理解をめざして（武田洋幸、岡本仁、成瀬清、堀寛　編）”．「蛋白質核酸酵素」12月増刊号, 45(17), 2943-2948 (2000)

4倍性魚類コイの *c-myc* 遺伝子2タイプの進化

二見邦彦・張　寰・岡本信明

生物の進化において，ゲノムの倍数性は極めて重要である。魚類における倍数性の解明は，脊椎動物の進化と遺伝子重複との関係を解く鍵であると考えられている。核局在性癌遺伝子 *c-myc* は，細胞周期を制御する最も重要な遺伝子の一つであり，4倍性魚類であるコイでは2タイプの *c-myc* 遺伝子が存在し，その両方が発現している。本稿では，倍数性進化の仮説に実証を与えるため，倍数化が比較的近い時期に起きたとされるコイの2タイプの *c-myc* 遺伝子に着目し，倍数化後の重複遺伝子の進化を機能との関連で紹介する。

Key words 【c-*myc*】【倍数性】【転写開始点】【Max】

Kunihiko Futami, Huan Zhang, Nobuaki Okamoto, 東京水産大学水産学部資源育成学科 （〒108-8477 港区港南 4-5-7） [Department of Aquatic Biosciences, Faculty of Fisheries, Tokyo University of Fisheries, Konan 4, Minato-ku, Tokyo 108-8477, Japan] E-mail: nokamoto@tokyo-u-fish.ac.jp

Molecular Evolution of Two c-myc Genes of a Tetraploid Teleost, The Common Carp

はじめに　Ohno らは，生物の進化において，遺伝子重複はきわめて重要であり，多細胞生物の多様化，組織の複雑化に密接にかかわっているという仮説を，1970 年代に提唱した[1)]。遺伝子重複には部分的重複と倍数性とがある。ひとつの遺伝子が重複すると，2 つになったうちのひとつは新しい機能を持つように進化する自由が与えられる。ゲノムの倍数性は，植物の進化では重要なはたらきを果たしてきたことが知られているが，脊椎動物においても約 5 億年前のカンブリア期におけるその爆発的進化（“カンブリア爆発”，“進化のビッグバン”などともよばれている）の主役であったと考えられている。倍数性化は染色体上のすべての遺伝子座が重複するため，構造遺伝子と調節遺伝子のバランスは保たれる。そして調節遺伝子の分岐により，構造遺伝子の発現の組織特異性が向上する。しかし，部分的重複では，必ずしも常に調節遺伝子が含まれるとは限らない。

魚類における倍数性の解明は，脊椎動物の進化と遺伝子重複との関係の謎を解く鍵であると考えられている。ヒトを始め，鳥類，爬虫類，両生類などで倍数性は知られているが，致死作用があったり，健康的でないものが多い。それに対し，魚類は現在でも倍数性化の能力を保持しており，異なる属間個体も容易に交配でき，染色体の可塑性を保持している[2,3)]。さらに，高等脊椎動物での倍数性化はかなり以前（数億年前）に数回にわたって起きたもので，重複した

遺伝子間の相同性が低くなり，いわゆる4倍性の2倍性化が進んでいる。そのため，重複した遺伝子の進化を解明することは困難である。しかし，魚類におけるゲノムの倍数性化は比較的最近（数千万～1 億年前）起きたため，二つの遺伝子間の相同性はまだ高い。したがって，ゲノムの倍数性化後の遺伝子の進化を研究する上で，倍数化魚類は格好のモデルになるといえる。

魚類において，サケ科やコイ科で多くの倍数性の例が報告されている。筆者らはこれまでに，4倍性魚類であるコイから2タイプの c-*myc* 遺伝子をクローニングし，さらに，その両方が発現していることを明らかにした。癌遺伝子としても知られている c-*myc* は，個体の発生・分化，および恒常性の維持にかかわる制御因子であるため，ヒトおよび高等脊椎動物で見出されたこの遺伝子に対応する遺伝子が魚類でも見いだされることは，当然のことといえる。本稿では，コイの重複した c-*myc* 遺伝子に焦点を合わせ，ゲノムの倍数性化による脊椎動物の進化を機能との関連で考察する。

Ⅰ. c-*myc* 遺伝子

核局在性癌遺伝子 *myc* は最初，MC29 という複製不能なトリレトロウイルスで同定され，v-*myc* とよばれた。MC29 ウイルスは，骨髄細胞腫，上皮性の悪

性腫瘍，肉腫，およびリンパ腫をひき起こす[4]。c-*myc*は v-*myc*の細胞性ホモログである。ヒトの染色体上には，c-*myc* 遺伝子と共通の構造的，機能的特徴を持つ遺伝子がいくつか存在することが報告されている[5]。それらは c-*myc*, N-*myc*, L-*myc*の 3 種類の重要な遺伝子の他に，B-*myc*, P-*myc*, R-*myc*, S-*myc*, N-*myc*2，および *myc*L2 といった計 9 種類のメンバーによる *myc* ファミリーを構成している。各遺伝子の構造は似ているものの，塩基配列およびアミノ酸配列の相同性は 50%以下である。これらの起源，類縁関係はいまだに不明である。

myc ファミリーのうち，c-*myc*, N-*myc*, および L-*myc* の 3 種類の主要な遺伝子は，ヒトの癌で最も頻繁に活性化のみられる癌遺伝子である。特に c-*myc* はヒト癌の約 1/3 で変異を起こしており[4]，*ras* などとならび現在最もよく研究されている癌遺伝子のひとつである。しかし，魚類においては，分子レベルでの腫瘍研究は必ずしも期待通りに進んでいるとはいいがたい。

哺乳類における c-*myc* 遺伝子の構造解析によれば，3 つのエキソンと 2 つのイントロンから成り立っていることが知られている[6]。c-*myc* 遺伝子の発現調節機構については，哺乳類ではよく研究されており，非コード領域である第 1 エキソンがその制御をしていることが明らかにされている。転写開始点は 4 ヵ所（P0，P1，P2 および P3）存在し，主要な転写開始点は 2 カ所（P1 と P2）

であるが，いずれの臓器，細胞種においてもP2からの転写産物のほうが多い。魚類では c-*myc* 遺伝子がニジマス，ゼブラフィッシュなどですでにクローニングされている[7]。ニジマスでは，第2イントロンと第2，第3エキソンがすでに明らかにされており，第3エキソンがよく保存されているほか，第2エキソンの中のボックス A と B の領域も特によく保存されている。

c-*myc* は，生体内では通常，細胞分裂を促進し，細胞分化を抑制することにより，細胞周期を制御している。c-*myc* の機能にはいまだ不明な点も多く残されているが，c-*myc* 遺伝子が個体の恒常性の維持に重要な役割を果たしていることは明らかである。そのため，脊椎動物のみならず，棘皮動物[8]や原索動物であるナメクジウオ（二見ら，未発表）などにおいても同遺伝子はよく保存されている。

II. コイ ***c-myc*** 遺伝子のクローニングと構造解析

筆者らのグループは，ヒトの c-*myc* をプローブとして，コイの末梢血白血球 cDNA ライブラリーから c-*myc* cDNA をクローニングした。このコイの c-*myc* cDNA をプローブとしてゲノムライブラリーのスクリーニングを行なったところ，2タイプの c-*myc*（*CAM1* と *CAM2*）がクローニングされた（図1)。ゲノムサザン解析の結果，コイの染色体上には2種類の c-*myc* の配列が存在するこ

とが明らかとなり，ノーザンブロット解析においても約 2.1 kb と約 1.5 kb の 2 種類の mRNA が確認された。*CAM*1 と *CAM*2 のアミノ酸配列は他種の c-Myc と高い相同性をもち（図 2），それぞれゼブラフィッシュで 90.4%と 90.9%，ニジマスで 72.9%と 72.8%，アフリカツメガエルで 59.2%と 57.4%，ニワトリで 57.1%と 57.6%，ヒトで 55.3%と 56.7%であった。倍数化によって生じた 2 タイプの c-*myc* 遺伝子である *CAM*1 と *CAM*2 の間の相同性は 94.2%であり，その値からコイにおける 4 倍性化は約 5,800 万年前に起こったと推定された[9]。

III．コイ c-*myc* 遺伝子の第 1 エキソンおよび転写開始点

c-*myc* 遺伝子の第 1 エキソンは非コード領域のエキソンで，哺乳類の場合，第 1 エキソンは c-*myc* 遺伝子の発現を制御しており，第 2 エキソンや第 3 エキソンと比べて，その進化速度は速いことが知られている。しかし，魚類においては，これまでに完全長の cDNA クローンが得られていなかったため，転写開始点はおろか第 1 エキソン領域の存在も明らかとなっていなかった。ヒトの c-*myc* 第 1 エキソンをプローブとして，コイのゲノム DNA に対し，サザンハイブリダイゼーションを行なっても，陽性シグナルが得られないことから，コイの c-*myc* には第 1 エキソンが存在しないか，あるいはヒトとの相同性がまっ

たくないか，このいずれかであると考えられた。遺伝子の転写開始部位の情報を得ることは，遺伝子発現制御の理解といった遺伝子機能の解析に直結している。そこで筆者らは，理論的には真の転写開始点を決定できるオリゴキャッピング法[10]の技術を用い，コイの c-*myc* 遺伝子から複数（heterogeneous）な転写開始点を見いだし（図 3)，魚類の c-*myc* 遺伝子に第 1 エキソンが存在することを初めて明らかにした[11]。

しかし， BLAST を用いてデータベースに登録されている配列との相同性検索を行なったところ，第 1 イントロン，第 1 エキソン，およびその 5'上流領域の配列は，ヒトを含め既知の配列との相同性はまったくなかった。さらに，推定されるプロモーター領域には，TATA ボックスや CCAAT ボックスの配列も存在しなかった。このことから，魚類の c-*myc* 遺伝子の発現は，高等脊椎動物の c-*myc* とは異なる方式で制御されているものと思われた。

また，*CAM*1 と *CAM*2 を比較したところ，第 1 エキソン領域は第 2 エキソン，第 3 エキソンと比べて両者の間の相同性が低く，転写開始点の位置や数も著しく異なっていた。したがって，*CAM*1 と *CAM*2 の間にも何らかの機能の分化が起きている可能性が示唆された。そこで，RT-PCR/サザンハイブリダイゼーション法により，2 タイプの c-*myc* 遺伝子の臓器別発現量を定量したところ（図 4)，この両者の間で発現パターンが異なり，転写レベルでの機能の分化が明ら

かとなった（二見ら，投稿準備中）。

Ⅳ. *c-myc* 関連遺伝子 *max* の発現

Myc ファミリータンパク質は，C 末端領域に塩基性領域－ヘリックス・ループ・ヘリックスーロイシンジッパードメイン（bHLH-LZ）といった，転写因子によくみられる構造をもっており，パートナーの Max（*myc* associate protein X）と結合し，ヘテロ二量体を形成する[12,13]。それによって DNA 上の特定の配列に結合し，細胞の増殖や分化にかかわる遺伝子の転写を活性化する。Max のホモ 2 量体も Myc/Max と同様の塩基配列に結合するが，標的遺伝子の転写を活性化する能力はなく，Myc/Max と拮抗し転写抑制にはたらく。すなわち，この両者の量的バランスにより細胞増殖が制御されているものと考えられている。

ゼブラフィッシュでは，*myc* ファミリー遺伝子が，Myc 蛋白質とヘテロ 2 量体をつくる *max* 遺伝子の発現と共同歩調をとっていることが明らかにされている[14]。そこで筆者らはまず，コイの *max* の cDNA 全長を RACE 法によりクローニングし，それをプローブとしたノーザンブロット解析を行なった（図 5）。卵巣における *max* の強い発現は，胚発生の初期に発現する L-*myc* との関係が予想されたものの，その発現パターンは，図 4 の *CAM2* の発現パターンと一致したが，*CAM1* とは明らかに異なった（二見ら，投稿準備中）。

また，近隣結合法により，c-Myc の分子系統樹を作成し，系統樹の枝の長さから分子進化速度を推定すると，倍数化後の *CAM1* の進化速度は *CAM2* に比べて約 1.6 倍速かった。これらのことから，*CAM2* は本来の c-*myc* としての役割を保持しており，魚類でも細胞増殖と細胞分化の制御にかかわっていることが示され，それに対し，*CAM1* は本来の c-*myc* とは異なる別の新しい機能をもつように進化しつつある状態であると考えられた。今後，蛋白質レベル，細胞レベルでの機能解析により，2 タイプの c-*myc* 遺伝子の役割を知ることができれば,それは倍数性進化の仮説に実証を与えることになるものと確信している。筆者らは現在，2 タイプの c-*myc* 遺伝子がどのような生理現象にかかわっているかを知るために,アンチセンス法による loss of function の実験系を構築中である。

おわりに　一般に，重複によって生じた遺伝子のうち，あるものは新しい機能を獲得し，またあるものは機能を失い，偽遺伝子となる。コイにおいても例外ではなく，現在，コイの重複した構造遺伝子の半分が機能を失ったと推定されている [15)]。たとえば *max* についても，筆者らの解析では，現在のところ 1 タイプしか発現が確認できていない。しかし，コイの 2 タイプの c-*myc* 遺伝子はともに発現しており，今でも機能を失っていないと思われる。c-*myc* は，遺伝

子の発現を調節する転写制御因子であり，哺乳類ではこの遺伝子の過剰発現は細胞の癌化に結びつく。ゲノムの倍数化により重複した構造遺伝子の半分が機能を失ったのに，なぜ重複したコイの c-*myc* はまだ両方が機能しているのであろうか？ コイの 2 タイプの c-*myc* 遺伝子は，それぞれ進化速度が異なり，タイプ 1 の c-*myc* 遺伝子はタイプ 2 より速く進化している。このことは，魚類においても，哺乳類にみられるような *myc* ファミリーが形成されつつあることを想定させる。したがって，魚類における c-*myc* 遺伝子の研究は，倍数性進化の仮説の実証に貢献するだけでなく，*myc* ファミリーの起源を知ることにより，*myc* の機能を進化の側面から明らかにすることができるかもしれない。

文 献

1) Ohno, S.: Evolution by Gene Duplication. P.160, Springer Verlag Press, Heidelberg, New York, (1970); 山岸秀夫・梁永 弘訳：遺伝子重複による進化，p.232，岩波書店 (1977)

2) 小島吉雄：魚類細胞遺伝学，pp.50-64，水交社，東京，(1983)

3) Turner, B. J.: Evolutionary Genetics of fishes, pp1-53, Plenum Press, New York and London, (1984)

4) Cole, M. D.: *Annu. Rev. Gen.* **20**, 361-384 (1986)

5) 口野嘉幸：*蛋白質核酸酵素*，**37**，865-867 (1992)

6) Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenior, G. and Leder, P.: *Cell*, **34**, 779-787 (1983)

7) 岡本信明・正仁親王：魚類の DNA（青木・隆島・平野編），pp. 414-423，恒星社厚生閣 (1997)

8) Walker, C. W., Boom, J. D. G. and Marsh, A. G.: *Oncogene,* **7**, 2007-2012 (1992)

9) Zhang, H., Okamoto, N. and Ikeda, Y.: *Gene*, **153**, 231-236 (1995)

10) Maruyama, K. and Sugano, S.: *Gene*, **138**, 171-174 (1994)

11) Futami, K., Komiya, T., Zhang, H. and Okamoto, N.: *Gene*, **245**, 43-47 (2000)

12) Blackwell, T. K., Kretzner, L., Blackwood, E.M., Eisenman R. N. and Weintraub H.: *Science*, **250**, 1149-1151 (1990)

13) Blackwood E. M. and Eizenman R. N.: *Science*, **251**, 1211-1217 (1991)

14) Schreiber-Agus, N., Horner, J., Torres, R., Fung-Chow Chiu and Depinho, R. A.: *Mol. Cell. Biol.*, **13**, 2765-2775 (1993)

15) Ferris, S. D., Whitt, G. S.: *Experientia*, **33**, 1299-1301 (1977)

図説明

図 1　コイの **2** タイプの ***c-myc*** 遺伝子（***CAM1*** と ***CAM2***）の制限酵素地図

Ba: *Bam*HI, Bg: *Bgl*II, E: *Eco*RI, H: *Hin*dIII, K: *Kpn*I, M: *Mbo*I, P: *Pst*I, Sa: *Sac*I, Sp: *Sph*I, X: *Xba*I.

図 2　コイの ***c-myc*** 遺伝子（***CAM1*** と ***CAM2***）と他種の ***c-myc*** 遺伝子とのアミノ酸配列の比較

ZEM : ゼブラフィッシュ, RTM : ニジマス, XLM : アフリカツメガエル, CHM : ニワトリ, HUM : ヒト。矢印は第 2 エキソンと第 3 エキソンの境界を示す。ボックスで囲まれた領域はよく保存されていることを示している。

図3　オリゴキャップ法により決定したコイ ***c-myc*** 遺伝子の **5'**末端部分の構造

clone1-18 は *CAM*1，19-21 は *CAM*2 にそれぞれ由来している。上に示した配列はゲノムの配列である。また、ここでは翻訳開始コドン（ATG）の A を+1 とした。

図4　**RT-PCR/**サザンハイブリダイゼーションによるコイ ***c-myc*** 遺伝子の発現量の定量

最下段は、内部標準として用いたβーアクチン

図 5　ノーザンブロット法により検出した各臓器におけるコイの *max* 遺伝子の発現

下段は、内部標準として用いたβーアクチン

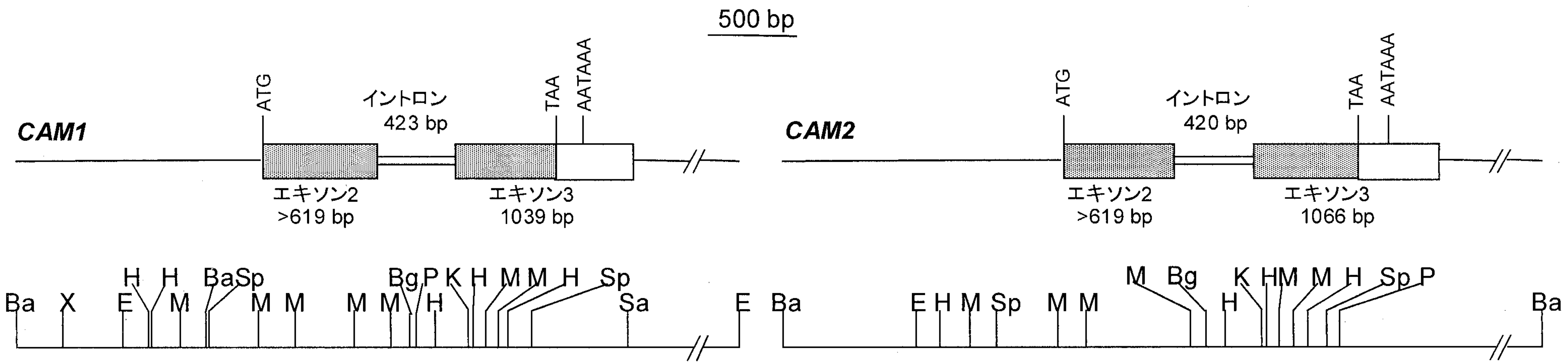


図 1（二見）

```
CAM1 MPVSASLAYKNYDYDYDSIQPYFYFDNDDEDFYHH-QQ--GQTQPPAPSEDIWKKFELLPTPPLSPSRRQSL--------
CAM2 ....S.............................-..--......................................--------
ZEM  .......C...........................-..--.....S.................................--------
RTM     NS...S...............V..E......--..-P..L.......................P-..SS--IFP-
XLM  ..LN.NFPS........-L..C.F.LEE-.N..--..--SRL........................S.QSS--LFP-
CHM  ..L....PS.........V......EEEE.N..LAA..RGSEL.......................S..AAASCFP-
HUM  ..LNV.FTNR...L....V.....C.EE-.N..QQQ..--SEL.......................SG.CSPSYVAV

CAM1 ----------------STAEQLEMVSEFLGDDVVNQSFIC-DADYSQSFIKSIIIQDCMWSGFSAAAKLEKVVSERLASL
CAM2 ----------------........................-.......................................K....
ZEM  ----------------.....................S....D.....................................K....
RTM  ----------------...D.....T.............-......T.L..............T...............
XLM  ----------------...D.....T....G.M.......-E..DEA-LL...V..........................K...Y
CHM  ----------------...D.....T.L..G.M.......-.P.DES-..V.............................K..TY
HUM  TPFSLRGDNDGGGGSF...D.....T.L..G.M.......-.P.DET-...N..................---...K...Y

CAM1 HAARKELMSDSSSN-------------------------RLNASYLQDVSTSASECIDP----------SVVFPYPLPES
CAM2 .......I......-------------------------.........L........G.----------........T..
ZEM  ..E....I...N..-------------------------.........L........G.----------........T.C
RTM  QT...DSAVGDNAECPT----------------------....N....PN.......G.NTSASECIGP.......IT.T
XLM  Q.S...SALS...PCQSQPPPSPLKSPSCHGSLSLGGTH.SSHGF...P.---.D.V..----------........NDT
CHM  Q.S.R.GGPAAA.RPGPPPSGPPPPPAGPAASAG---------L..H.LGAA.AD....----------........S.R
HUM  Q....DSG.PNPARGHSVCSTSS-------------------L....LGAA.......----------........ND.

CAM1 GKSSKVAPSEPM-----------------------PVL--DTPPN--SSSSSGSDSEEEEEEEEEEEEEEEEE----EID
CAM2 S...........-----------------------L..--.....--..........D.................EEEE...
ZEM  ..AG...SPQ..-----------------------L..--H....--..........D....D..........EEEE...
RTM  P.P.....PTD------------------------LA.--......SG..........DDD...DD.D...--------...
XLM  ISNASSPCQD-------------------------LI.--E...I--..N..S.E....P.D.D.DCD...-------...
CHM  APRAAPPGAN.ALL---------------------LGV--.....------TTS......Q.DD.--------------...
HUM  SSPKSC.SQDSSAFSPSSDSLLSSTESSPQGSPEPL..HEE...------TTS......Q.D..--------------...

CAM1 VVTVEKRQKKNETAVSDSRYP--------SPLVLKRCHVSTHQHNYAAHPSTRHDQPAVKRLRLEAS-SN-SNS-----R
CAM2 .........R..AE.......--------.................................S.S..N.S.N----.
ZEM  .........RH..DA.E....--------...................................NNHSI..SSSSN.
RTM  ........AVKRCDP.T.ETR------HH......................E.........N.S.R-VLK-----Q
XLM  ........SASKRVE.S.HSQ--PSRPHY..........PI.......S...KV.YVSS..AK..-.NIR-VLK-----Q
CHM  ...LAEANESESSTE.STEASEEHCKPHH..........NI.......P...KVEY..A...K.D-.-GR-VLK-----Q
HUM  ..S.....APGKRSE.G.PSAGGHSKPPH...................P....K.Y..A..VK.D-.-VR-VLR-----Q

CAM1 HVKQRKCTSPRTSDSEDNDKRRTHNVLERQRRNELKLSFFALRDEIPDVANNEKAAKVVILKKATECIHSMQLDEQRLLS
CAM2 QG...........................................E................................
ZEM  .......A.....................................E................................
RTM  ISSN...S.......T..Y..................................................Y...T.....VN
XLM  ISNN...A...S....E....K.......................QV.E..S....P..........YAI.L.E..R..IR
CHM  ISNN...S.........E...........................Q..E.......P..........YVL.I.S..H..IA
HUM  ISNN........S..T.E.V.................R.......Q..........P.........AY.L.V.AE..K.I.

CAM1 IKEQLRRKSEQLKHRLQLLRSSH*
CAM2 .................Q.....*
ZEM  .................R.....*
RTM  L.........H..QK.AQ.QN.CLSSKRH*
XLM  ET...KYRK....Q...Q..NFV*
CHM  E......RR.....K.EQ..N.RA*
HUM  EEDL..KRR.....K.EQ..N.CA*
```

図 2（二見）

```
                    -569                                                                                        -474
CAM1        ------ACTGACACCCAGGCTATCAGTCTGTCCAGCACCTGAGGGGTGCTCGAGTGCGTACGTCGATTACGGTTTCCATCACTTTCTATTTGGCCAGGACTC
CAM2        ------T...T...A...A..........T.....T.....A...A.ATAG.A...AG...T.....AT..............C.....A.........
                    -596          -581                                                                          -501
クローン 1                  AGACTATCAGTCTGTTCA------
         2                       ATCAGTCTGTTCAGCATC------
         3                       ATCAGTCTGTTCAGCATC------
         4                       ATCAGTCTGTTCAGCATC------
         5                       ATCAGTCTGTTCAGCATC------
         6                       ATCAGTCTGTTCAGCATC------
         7                       ATCAGTCTGTTCAGCATC------
         8                         CAGTCTGTTCAGCATCTG------
         9                         CAGTCTGTTCAGCATCTG------
        10                                   AGCATCTGAGAGGTACAT------
        11                                                            GCGTTGATTACATTTTCC------
        12                                                                              ATCACTTTCCATTTGACC------
        13                                                                                      CCATTTGACCAGGA---
          -473                                                                                              -375
CAM1        TGATTCCACTTATGCTGCAAAAGACCGGAGTAAACTATTGGATCTTACGTTTTGTTCCTTATTTTTTGCTCTCTTTTCATTGCATTGCGTCTCGAACGA---
CAM2        ..............T.....G...............G..........T.....C....G........T..........C.....C.........GA..---
          -500                                                                                              -402
クローン 14         CACTTATGTTGCAAGAGA------
        15           CTTATGTTGCAAGAGACC------
        16                   GCAAGAGACCGGAGTAAA------
        17                                                                                         GCATCGCGTCTCGAGA---
        18                                                                                             CGTCTCGAGAGA---
        19           CTTATGCTGCAAAAGACC------
        20                   GCAAAAGACCGGAGTAAA------
        21                                                                                                   GAACGA---
```

図 3（二見）

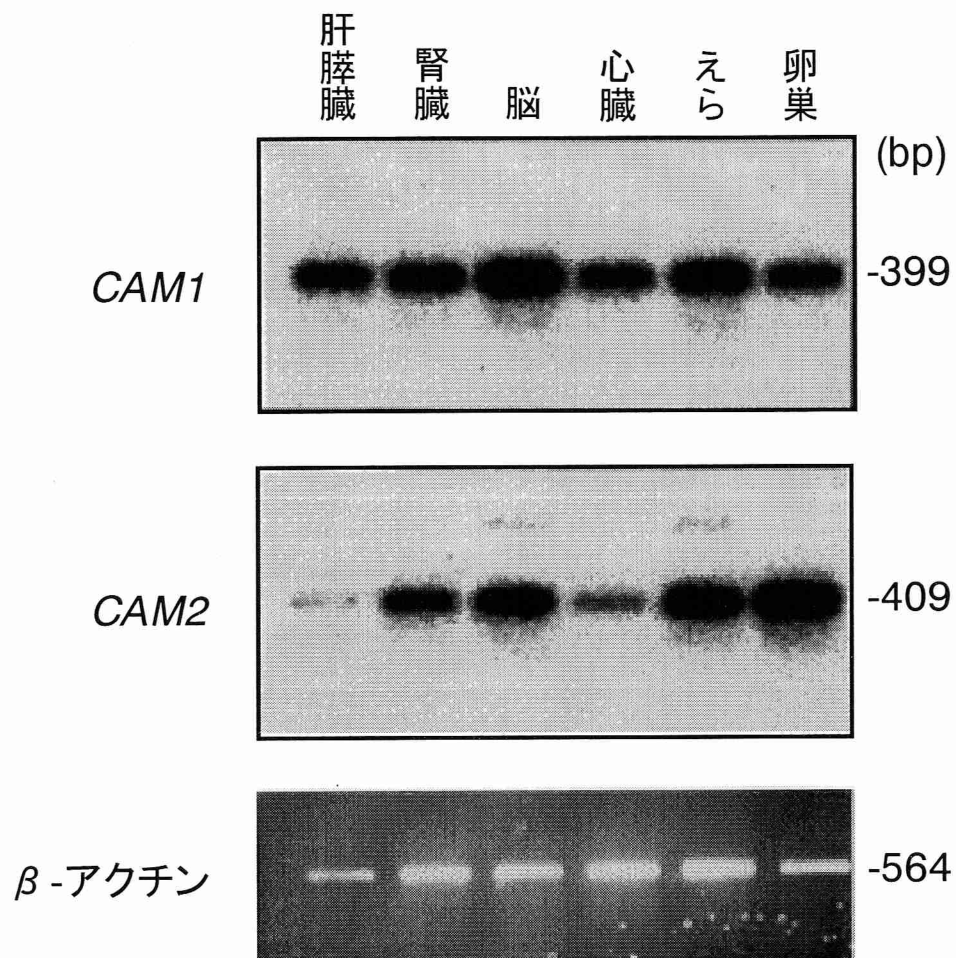


図4（二見）

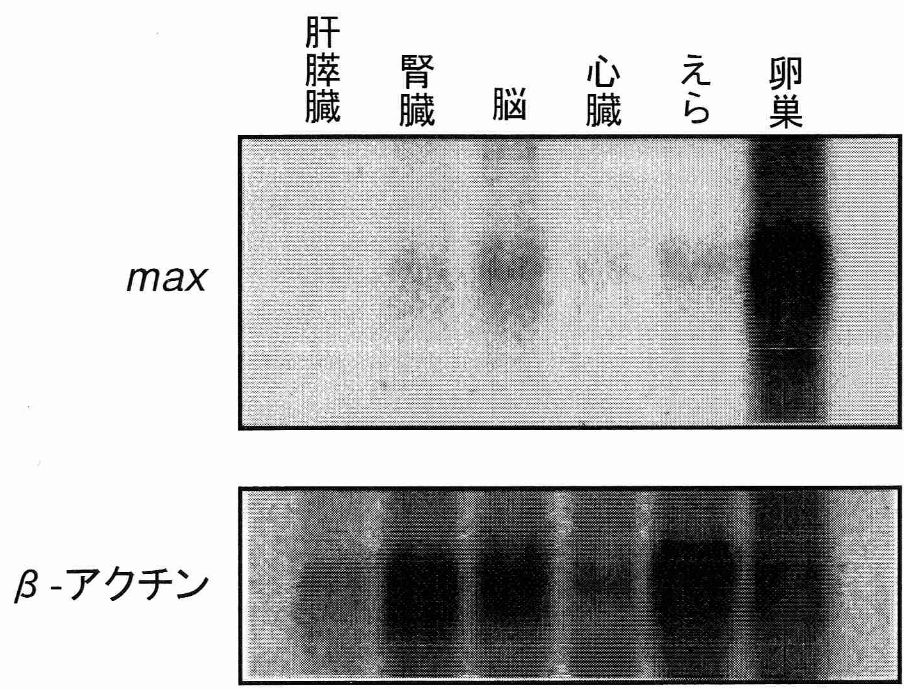


図 5（二見）