

メダカゲノム情報によるステロール生合成遺伝子の解析 ～メダカ *DHCR7* 遺伝子全構造の推定～

Analysis of sterol biosynthesis genes
from the genome information of Medaka *Oryzias latipes*
The estimation of complete gene structure of *DHCR7*

西野 秀昭

甲斐 慎

Hideaki NISHINO

Makoto KAI

福岡教育大学・理科教育講座 (生物)

(平成25年 9月30日受理)

Abstract

The possible complete gene structure of Medaka *Oryzias latipes* *DHCR7*, in which the enzyme sterol $\Delta 7$ -reductase is encoded, was investigated through BLAST search by comparison of National BioResource Project, Medaka genome database, with the cDNA base sequences of *Homo sapiens* and *Danio rerio* in order to elucidate the mechanism of biological evolution of sterol biosynthesis genes. As the results, all possible exon structures, 1st exon to 7th exon consisting of 1419 bp open reading frame containing stop codon and the estimated 472 amino acid residues, of *DHCR7* of Medaka *Oryzias latipes* were extracted as the strong candidate in the Medaka third chromosome database. The estimated primary structures of the all exons except first exon were very similar between Medaka *Oryzias latipes*, *Danio rerio*, and *Homo sapiens*. The Medaka *DHCR7* also contains S-1 sequence, which is a unique motif in the primary structure of sterol $\Delta 7$ -reductase.

要 約

ステロール生合成遺伝子が、生物進化の過程で創生されてきた機構を探究するため、ステロール生合成経路の酵素 sterol $\Delta 7$ -reductase をコードする遺伝子 *DHCR7* の構造を複数の生物で比較する目的で、メダカ *Oryzias latipes* の *DHCR7* 構造の全解明に取り組んだ。既に推定されている *DHCR7* の exon2 と 3 の候補配列で National BioResource Project Medaka Genome Database を対象に BLAST 検索を行った。その結果、第3染色体で、メダカ *Oryzias latipes* の *DHCR7* の第1から第7 exon の候補全てを見出すことができた。開始コドンから終止コドンまでの読み枠の塩基対数は1419個、推定アミノ酸残基数は472個であった。7個という exon 数は、ヒト *Homo sapiens* やゼブラフィッシュ *Danio rerio* と同じであった。推定アミノ酸配列も第1 exon 以外は三者で高度に類似し、sterol $\Delta 7$ -reductase のコンセンサス配列 S-1 も有していた。これらから、Medaka Genome Database 第3染色体から得られた塩基配列は、メダカ *DHCR7* 遺伝子であると考えられた。

1. はじめに

ステロール生合成遺伝子は、真核生物にのみ見出され、前核単細胞から真核多細胞へと生物が進化する過程を考察する一つの有力な手段と考えられる¹⁾。ステロール構造体によって、細胞膜はどのような変形にも耐えられるという意味での強靱性を増すとともに、このような細胞膜の柔軟性によって、細胞の大型化や貧食能の獲得へとつながったと推定されている²⁾。

メダカ (*Oryzias latipes*) は、日本で独自に開発された脊椎動物のモデル生物で、生物学や医学のみならず、環境科学、水産学などでも広く利用され、英語名も Medaka と表記される^{3) 4)}。さらに、比較ゲノム生物学の進展に伴い、メダカとヒト (*Homo sapiens*) のゲノム遺伝子構造には多くの共通性が見出されている⁴⁾。従って例えば、メダカの遺伝的変異体を解析することで、ヒト遺伝病の原因を解明することも可能となることが考えられる。メダカは胚発生に関わる 300 余の突然変異体が知られ⁴⁾、同様にモデル生物であるゼブラフィッシュ (*Danio rerio*) など他の硬骨魚類の突然変異体では見られないものが多く、研究対象としての貴重なコレクションとされている。またメダカは日本の自生種であり、野生集団の遺伝子解析を行うことによって自然環境の成り立ちや、その保全に新しい視点を与える事も考えられている。

ステロール生合成経路の最終段階でステロイド骨格 7 位二重結合を還元する酵素をコードしている遺伝子 *DHCR7* は、高等動物の正常な形態形成にも重要なはたらきを担うことが知られている³⁾。本研究では、メダカゲノムデータベースの情報や GenBank 情報を活用し、未だ一部しか構造が知られていない³⁾ メダカ *DHCR7* 遺伝子の全構造の推定を行った。アメリカの国立衛生研究所 National Institute of Health (NIH) により運営されている web サイト、National Center for Biotechnology Information (NCBI)⁵⁾ の GenBank という世界的な DNA データベース〔日本の DNA Data Bank of Japan (DDBJ)、欧州の EBI/EMBL も相互にデータを共有している〕にはメダカ *DHCR7* の情報は未だ見られないが、日本独自の database である National BioResource Project (NBRP) によるメダカゲノムプロジェクトでは、ニホンメダカ *Oryzias latipes* のゲノム解析によって、約 100 万リード、726Mbp の情報が既に塩基配列の決定がなされ、データベース化さ

れている⁴⁾。メダカゲノムサイズは約 800Mbp と推定されているので、全ゲノムの 0.9 × 分に相当する⁴⁾。

一方でゼブラフィッシュは、脊椎動物の小型魚類モデル生物として多くの genome database や mRNA 配列データが NCBI に登録されている。ヒト *DHCR7* 遺伝子構造及び cDNA のデータも NCBI に登録されているので、ここからゼブラフィッシュとヒトの *DHCR7* 遺伝子構造の比較を行い、類似性が高い構造は、メダカゲノムの *DHCR7* でも同様に類似していると考えられることができる。このような方針で西山ら³⁾ は Medaka genome database を検索したが、exon2 及び exon3 を含む genome 塩基配列までしか検出できなかった。そこで本研究では、NBRP の Medaka genome database のうち、Ensembl NBRP サイトを利用して染色体情報の方から、既に見出されている exon2 及び exon3 の塩基配列を含む染色体を見出すことによって、メダカの *DHCR7* 全構造を解析することとした。

さらに本研究は、平成 20 年 9 月改訂の中学校学習指導要領・理科編に新たに導入された理科の単元、「生物の変遷と進化」及び「遺伝の規則性と遺伝子」⁶⁾ の観察・実験を含む授業に利用できる生物教材の開発も兼ねている。「生物の変遷と進化」では、生徒達が小学校から馴染んでいる教材であるメダカを使って、同じ働きを持つ遺伝子であれば種類が違ってても良く似た構造であることを通じて生物進化の理解に使える可能性がある。「遺伝の規則性と遺伝子」単元では「内容の取り扱い」において、「…また、遺伝子に変化が起きて形質が変化することがあることや遺伝子の本体が DNA であることにも触れること。」とあり、この単元の観察・実験で利用できる生物教材の開発が待たれている。*DHCR7* 遺伝子の変異の種類によっては、ヒトにおいては単眼症や前脳の未発達など形態形成異常などを引き起こす。生徒達に馴染みがあるメダカでも、受精卵の胚への初期発生段階で単眼になる変異が NBRP の Medaka bioresource に見出されることが知られている⁴⁾。その親は原因遺伝子についてヘテロ接合体とされ、一つは正常な遺伝子を持つが、ヘテロ接合体の親から 4 分の 1 の確率で生まれる劣性ホモ接合体では単眼になり致死になるものと考えられている。その原因遺伝子については *DHCR7* の変異が疑われているが、*DHCR7* 遺伝子の全構造が未解明の為、未だ特定されていない。本研究は、メダカの *DHCR7* 全構造解明がメダカの単眼の発

生メカニズム解明に役立てるとともに、遺伝子の本体がDNAであることを理科授業の観察・実験を通じて理解できる生物教材開発をも目的にしている。

2. 研究方法

配列変換フリーソフト SeqConv v0.81 を用いて、メダカ *DHCR7* の exon2 と exon3 候補の塩基配列³⁾ をアミノ酸配列へと置換した。その推定アミノ酸配列を用いて、NBRP Medaka Hd-rR genome database を対象に、アミノ酸配列から塩基配列を検索するため、tblastn による検索³⁾ を行った。NBRP Medaka 内の Medaka Map をクリックし、Ensembl NBRP を開いた。その Medaka をクリックし、Ensembl NBRP Medaka を開いた。その Search Ensembl NBRP *Olyzias latipes* に「*DHCR7*」と入力し、Ensembl NBRP Medaka Unisearch を開いた。ENSORL00000019316 をクリックし、Ensembl NBRP Medaka TransView を開き、Genomic sequence をクリックした。得られた塩基配列の中から exon2 と exon3 に加えて他の exon 候補を、GU と AG に挟まれた intron 及び、ヒトとゼブラフィッシュの遺伝子構造に注意しながら見出した。ヒト *Homo sapiens DHCR7* の cDNA 情報は、accession number が NM_001360 というデータに、また、*Danio rerio DHCR7* の cDNA 情報は、accession number NM_201330 に登録されているものを用いた。cDNA (mRNA) と genome DNA の塩基配列を比較し、GU-AG 則³⁾ を用いて、exon 構造を塩基配列レベルで見出した。ヒト genome における *DHCR7* の塩基配列は、accession number が AF110060 の遺伝子情報を用いた。*Danio rerio* のゲノム遺伝子情報は、NC_007113 のものを用いた。遺伝子構造はいずれも、exon は 7 個で、読み枠はヒトが 475 個、ゼブラフィッシュが 479 個のアミノ酸残基からなっていた。またアミノ酸配列は高度に一致し、exon1 の C 末端側アミノ酸配列 QWGRAW は完全に一致していた（ただし最後のトリプトファン W のコドン TGG は最初の TG のみ exon1 に存在し、最後の G は exon2 の最初の塩基だった）。しかし exon1 はそれ以外はアミノ酸残基数も含め、ほとんど類似性が見いだされなかった。

3. 結果と考察

NBRP Medaka 内の Medaka Map から、Ensembl NBRP へ飛び、その Medaka から、Ensembl NBRP Medaka を開いた。その Search Ensembl NBRP *Olyzias latipes* に「*DHCR7*」と入力し、Ensembl NBRP Medaka Unisearch を開いた。得られた情報 ENSORL00000019316 をクリックし、Ensembl NBRP Medaka TransView を開き、Genomic sequence をクリックした。得られた塩基配列の中から exon2 と exon3 に加えて他の exon 候補を、GU と AG に挟まれた intron とヒトとゼブラフィッシュの遺伝子構造に注意しながら見出した（図 1）。その結果、第 3 染色体の 34524922 から 34533555 までに含まれる exon-intron 構造を見出した。既に genome database 上で exon 候補が 8 カ所ライトアップされていた（図 1 の黒に白抜き文字）。しかし最初の exon に開始コドンがなく、2 番目の exon 中の ATG を開始コドンとすると、ヒトやゼブラフィッシュの exon1 の C 末端側に見出されている共通アミノ酸配列 QWGRAW がメダカでも見出された（図 2）。このことから、本来の exon1 はこの開始コドンから始まり、上記共通配列 QWGRW の C 末端アミノ酸残基 W（トリプトファン）のコドン TGG の TG までが exon1 に含まれることが推定された。このように exon1 の C 末端アミノ酸残基 W のコドンが次の exon との間で分断されていることは、ヒトやゼブラフィッシュでも見出される（図 3）。既に西山ら³⁾ によって見出されている exon2 及び exon3 とともに、合計 7 つの exon を見出してアミノ酸配列を推定した（表 1）。その結果、ORF（読み枠）は終止コドンを含め 1419bp、コードされているアミノ酸残基は 472 個と推定された。

推定されたメダカ *DHCR7* 候補遺伝子を確認するため、ヒト及びゼブラフィッシュの *DHCR7* と一次構造の比較を行った（図 4）。その結果、exon1 のアミノ末端側のアミノ酸配列は比較したメダカ・ヒト・ゼブラフィッシュでほとんど似ていなかったが、それ以外の配列はほとんど一致するか類似の性質を示すアミノ酸残基が見いだされた。全体では、同じ位置に同じアミノ酸残基なのが 67.2% で、類似の性質を示すアミノ酸残基まで含めると 77.1% と高度な類似性が見出された。この類似性の程度は、遺伝子創成の起源を探る手立てとして使えるかもしれない。

また、ヒトやゼブラフィッシュの *DHCR7* 同様、そのコードしている sterol $\Delta 7$ -reductase のコン

センサス配列である S-1 配列⁷⁾も存在していた(図4)。今後、ヒトなどで Lamin B receptor や sterol C14-reductase など、機能が異なるにも関わらず *DHCR7* と S-1 配列以外で高度に一次構造が類似していることが知られているタンパク質や塩基配列をメダカで比較し、メダカ *DHCR7* を含めた共通祖先遺伝子から S-1 配列の挿入による *DHCR7* 遺伝子の起源に関わる議論の材料としていかなければならない。

また、本研究で見出すことができたメダカ *Oryzias latipes DHCR7* の各 exon が実際に存在しているのか、また exon1 の推定が正しいのか、メダカからゲノム DNA を抽出し、推定した exon の最初と最後、また、前後の intron と推定される部分からプライマーをデザイン及び合成して、PCR (Polymerase Chain Reaction) によって増幅できるかどうか確認することも、酵母細胞などでの機能発現解析とともに今後の課題である。

これらの成果を受けて、メダカの生物教材化の課題も見出される。すなわち中学校理科「生物の変遷と進化」, 「遺伝の規則性と遺伝子」単元での観察・実験向けに、メダカをより利用可能な生物教材へと開発することが期待される。メダカは、NBRP の Bioresource を活用し、変異 *DHCR7* 遺伝子をヘテロ接合体で持つ親メダカの交配によって4個に1個の確率で得られる劣性ホモ接合体での単眼症を、胚発生段階で、中学校にも配備されている生物顕微鏡などで検出する手法の開発、中学校の理科室でもできる PCR の手法を開発し、これによって単眼症の胚から変異部位を検出する方法、外部発注を含め、塩基配列を簡易に決定する手続きの具体化や手法などが考えられ、これらの具体化・実現化が待たれる。

4. 参考文献

- 1) 西野秀昭・牧尾友美：生物界ステロール生合成遺伝子の分子進化に関する考察，福岡教育大学紀要，第58号第3分冊，75-88，平成21(2009)年2月
- 2) 田川邦夫：からだの生化学，初版第1刷，108-110 & 113-114，1993年4月30日，(株)宝酒造
- 3) 西山智子・清水誠之・西野秀昭：メダカゲノム情報によるステロール生合成遺伝子の解析～メダカ *DHCR7* の遺伝子構造～，福岡教育大学紀要，第61号，第3分冊，47-59，2012
- 4) ナショナル Bioresource プロジェクト (NBRP) 情報公開サイト (<http://www.nbrp.jp/index.jsp>)，メダカデータベース (<http://www.shigen.nig.ac.jp/medaka/>) Medaka Genome Database, http://mbase.bioweb.ne.jp/~dclust/medaka_top.html (2013年9月25日最終アクセス)
- 5) National Center for Biotechnology Information (NCBI) HP, <http://www.ncbi.nlm.nih.gov/> (2013年9月25日最終アクセス)
- 6) 文部科学省：中学校学習指導要領解説・理科編，77-78及び85，平成20年9月，大日本図書
- 7) Nishino, H., & Ishibashi, T.: Mammalian sterol $\Delta 7$ -reductase: cDNA cloning and proposition as a potential sterol-sensing membrane protein, *Biomedical Res.*, 20(5) 305-308, 1999

本研究は、科学研究費助成事業（学術研究助成基金助成金）（基盤研究（C）課題番号 24501104）の交付を受けて行った成果である。

THIS STYLE: Location of ENSORLG00000015429 exons THIS STYLE: Location of Ensembl exons
>chromosome:MEDAKA1:3:34524922:34533555:-1
ACAGTGTGGAAGAAGAAAATGAAAAAGTATTTGGAGGATTGCTTTTTAATTTTAAATTTTGTAGTCAAAAAGTGGCGCTTCATTTTTGAAAATGAAA
AAGCATTTTACAGTATTGAACTTTATTTTTAATTTATGCTACAAAAATGCTTCCATACATACCGGTTGCATTGCAATGTCTACCCACCTCTACAT
GGCAGTTATTCATTTGCGGATTTTATACAAAAAAAATAAAAAAAATACAAATGATCAAGAAAACGGGTGAGTATTTCTCTTCAAGCAGTC
CAGTATAACCAGTACTCTAAAATATCGCTCAGGTGTAATAATGAAAGGTGTACATGTTTTTAAACAGCTCCAAGTAGACGCACGCGGTTGT
TTTGTCTATCCAATCTGCTCGCATCAAGCGCTTGGAGTTTGTGAGCGCTGTGGGACAAAACCAATCAGAGGCTGAGAAATCCAGCACTGGCCC
GCCTGTCAATGAAATCATATGTTGTAATTTATCCAACCTAATGGTCTTATGGCGGGGCTATCTCCCTTACGCGGTTACAAAATAAGAGCCCTT
TTTCTGCCGTTCCACAGCGGTTGAAAGAGCTCCAGCTCCGGCTAAACTCACAGCTGCAGCGCGGTGAGTCCGTTTTTACGACCATGCAAGTCT
GAGTCTGCGCACTACGTATGATGCTGTTAAAAATCCTGTTTTAATGTAACCTTTGCTGTTATTTTAAACGCTCAAATGTTAACTTGGGGCAAAA
AAAGGAAGAAAAACGCATTTCCGCTTGTAAAGATGATAAGATCGAACAGATTTGTTTCAAAGAAAACCCCTCGATGTTCTGTCGGGTTACATAAT
AACGGTGTCTTCCCTCCTGTTACATATTTCTAAATTTCTCACGACACATTAACCTTAAAAAAGGCTAAATTTTGCAAAATAAACAGCATTTTTTA
GCTAATAAAGGAGTTTACTGTGAGGTGAGGAAGTGTGTAACACCAACGTTGACTGGTTTACATTGGCTGAGGTGAGCCCTATTTTGGCCCTT
TTCCCCCTGAAAACATACACCTAAATACCAGCAATATCTTAAGATGTTCATCAAAATTTGATGCTTAATTTTCATAATAATTTCTCGAATAC
ATTATGATGTAACAACTAAAACATATGACTTCCAGGTACCAATTTGACTTTTCTGCAATCCAGACAGAACCTGAATAAAGTTACACAAA
ACTTGTCCAGATTCATATGCTGTCTGATAAGATTAATTTCTGTTCTTCTTCTCAGCTTTTGGCTATCTCTGGGCTTCCAGCTTTTGGCTATCT
AGGCGCGCGCAGCAGCCTCCAATGGGAAAGCCCTGAGCAGCGCGCAGCAGTCCAGCACAGTGGGGCAGAGCATGCTCAGCACTGATGGCGACCT
TTCAAACAATGTGCATTTTGTGACTTAATATCGTCCAAATCATATTAATGGCGGTTATTTAAATGCATCCTTTATGGCACTGAAGACTAAAG
CACGTTACAGTCCACAATGATGTAAACAGGTTTCCATTCAAGATGCTGGTGTAGCGGTGAGCCAGAAGCCCTGGTTCAAATCCCTGTGACTTT
TTTTCTGATGTTGAAGTGTGCTGTTGCTCAATTTCTCAGCAATTTCACTGACATTTAAATACCTTAAATTTATCCCTTAAATGTTGATGTTT
TAGTGGGAATGTGAAAGTGAAGTGTGTTTTGTCTCTGCAATGGACTGGCTAACCCGGACAGTGGATGGGTTAGGCTCTAGCAGCTCCGTGACCCC
AAAAGGGATAAAGCGGGTTTAGAGAGTAGATGGAAGTTCATCATATAGTCTTCAATCCAGGAGCTCACAACATTTTGTCAATTTCTGATGTTT
ACACTAGGCTGCACGATGAGGAAATTTCAAAATGGCAGATTTAGTGAACAAATTTGTGATGACGATATAACTTGCAAAATATGAAACAAAT
AGCAAAACCAAAAACCACTGGTTGCTCAATTTCACTGACAGATTTAAATACCTTAAATTTATCCCTTAAATGTTGATGTTTCAACCTT
TTTTGTACCCGACAGCCGTACAACGCAAGACATGTTACCGTGGACCGGGGTTGATTTTTCTCTTTAGTAATAATAAACAATAATAAAT
GAACTAAATCCACCGTGAATCATGTGCAACTTTATTAGCGGTGCTCGTAACGTCACACCAACAACAAAACATCACAATGAGCAACATCTTT
TTCCACCAACCAATCGCACACTCTGGTTGCTATGGTAAAGTGTGACGGATGTAACAGAGAGAAATCCGGTCACTTTTCAAAGAAAACCTTTT
TTTTCGAGAAAAAAACCAATAAAAATCGAAATTTTCTTTCTGTCGGCTCCCGGTTGCCGTTGCCGCTGACCGGTTCCCGGACCGCCGTTG
TTGGGACCACTGCTCTAAATGACCCCTGCTCTACATGGGAGGAAAAATCATAAATCGTGTGGCGTCCATCTGATTACCCAAATGCATAAAGA
CATTTATTTGATTGTTAAGTACTTCAAGCTGCAATAAGATTGTTTTGCTTTTACACAGTCTTCTGTCGCTTTTGGCATGATGCTGTTATA
CAGTTTGATATGCGGATAATGATAAATATGCAATATTTGCTGGCCGTTTACATTGAGAGCTGCATAAGCGTTCCAGTGTTTTTCTTTT
AAATGATGTTGCGGATTTAGCCAAAATCCCAAAAATCTGTGTCGTTTTCTGGGACATAGTTTCTGAGAGACAGAAATTTCAATTTGAAAT
TTGGTCTTAAATAATGTGGTATTTCTTTGAACCTTTAACATTTTACAGAGAGATGTTATTTAAATGAATCCTTTACATGATTTCTATGATG
TTTAGTTATTTACCGTGAATTTCTGTTGAGGTAGTCACTGTTTTTACAGTACCGTGTACTCCTCTCACGTTACCCCCACGTTGACTCCTCC
AGGAGGTTGAGCTGGTTCTCTCATACCGGTGGTGGCTCCTCATCTTTGCTCCTTCAATCTGTTTACTTCTGTTGAGCCCTGTGACCACT
TCCAGTGTCCCTGACGAGCCCTTGTGCTGAGCTTTATTACGGGGACAGTGCAGCTTCTCTCCATCTGTCGCGCGGGCCCTCTTCTCTCCTGTC
GGCGGCTAAGATATACCGCTTTTGGGTGGGCTTCCAGTTGAGGCTGAGGGACGATGGCGCTGTGCAAGTTGCATATTTTAGGGGTTTTCA
CCGTCTGTTTGTCTGCTCCAGCTCTCTACCTGTGTGTCCTGATGTACCCACAAGTTTATTTCCAGGTTATGTTGGGGGGTGCAGGATGG
AGCACGAATCCTGCAGCTAACATCTGTTAATGCGGTTTTGTTATTTATGTTATTGCTACGACAATTCAGGTTATTTTCCAGATTTATCATA
AAAAAGGTTGTTGTTTCTCTCTCTGCTCCACTAGCTCTCATTAACAAGTATGATGATCAACGGGCTGCAGTGTGGCTGATCACGCACTGCCCT
GTCGTTCCCAACCCCGTACTTCCACTGGTTTTCTCCACCATCATCTTCGATAAATCTGATCCCGCTGATGTTCCGCTGCGCAATTTGTCTCGC
TACCGGTTGCCAGTTTGCCTTCTGCAAGGCTTACATCTTCCCACCAACCCCGAGGACTGTGAGCCCTGCACCGACCCGCCATCGGGACA
CCTCTAACACCCCTTAAACCCCTAAACCAATGCGCGTTGATTTACAGCAAGTTTACAGGAAATGCTTCTTACAACCTACATGATGGCATCCGAT
TTAACCCGCGCATCGGCAAGTGGTTCGACTTCAAGCTTTTCTTCAACGGCGGACCCGGTATTGTGGCCTGGACTCTAATCAATCTCTCTACAT
CTCCAAGCAGCAAGATTCTTGAACCCCTCACCAACTCATAGTTCTGCTGACCATGATTTCCAGCTCCGTCAGCTACGTCCTGTGGGACCGGACCTGA
CGTAATCTACTGTCCTGCTTCGGATTCAGGCGAGATCCTGTTAAGGGAGTCAATAGACTCCACTGGACAGGTAGTCTCACTCTGAGAGAAGGA
GAAGCAGGTAGAGGGGCGGGAGAGGGGGTTGGACAAAACACACAAGGAATGCTTCTTAATAAAGAAAGCGTTTTGTCATATCTGAAAAATGAAC
ATATTATATTGGGTTACTGGATTATTACTCTAGGTGAGGGGTGCAACTCATTTTGGTTTGGGGCCACATTTAATCCATCTGATCCCAAGT
GTCCGGACAGTGACATAAAAAGCAACATAACAGCTTTTGTGTTTTTCTTGAATGAGAAAAATACCTCAACCGCAAAAACACATACATTC
(2256bp deleted)
TCTCTCTTCTCCCTCTTCTCTTTCCGTCGGGTCCAACCAAAAGATTTTCAGACATGTTTGAATAATAAAGTTTGGCCCTCAATTACAAAA
GGGGTTTATTCAGACATACCTTTGGTTGTCTGAAGATTAATAACCCCTTGTAAAGTAAAATATGTCACCAACCAAGGGCCCTCAGCTCTC
ATCTGACTGCTTACGCTGTTGGACAGGACAAGTTAAGAAAGAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
TAATGAAAAGCAATCCACCAAAATGCTTTTTCAATTTCTTCTTCAACACAGCTTATTTCTGTCACCTAATTTAAATTTAAATGAAATGGTAGTT
GACATTTCAATTTTCAAACGTTCTCTGGTAAATGTGTAGCAGCGGTACCAGACAGCCCTACCTAACCCCACTTTGATTACCAACAGTTTGGAG
AAAGTGAACGCGAGGTTCTAACACCAATGTTTGGAGAAAATCATGCAAAACCATGACGAGCTCTAATGTTTCTGCTGTTGTTTACCCGTGA
CGGGATTTACGCTGCTGGATTTCTTCTGGAATGAGCGGTGTTACCTGAAAGCAATGACATCTGCCACGACCACTCCGGATGGTATCTGGGATGG
GGGACTGCGTGTGGCTGCCATACCTGTACAGCTGCAGCTACGTTGACGACGATCTCAGATGCTCCATAAGAGCCAAAACACATAATCGTTTATCA
GAAAGCGTAAATGAATAGCTTTTAACTCTATTAAAGTTTCAAACAGAACATGACACCCGTGAGACCTTGGACTCACAGCTGGAATGATATCA
GTATCCGTGATGGCGGTAGATTGAAGCTTATCTTCTGGACCTGCAGCTGGCTGTACCTGGTTTACCACCCCGTCCAGCTCTCCAGCGCTCC
GCCCTCGCAGTCTGCTGCTCGGCTCCTGGGATTTACATATTTCCGCTCCACCAACCAACCAAGGATCTCTTCCGCTCGCCACGATGGCTCCT
GCACAATTTGGGGCGCAAGCCGCTTTTACATCGAATGCTCGTACCCTCCGCGGATGGCGGCTTGCACCGAAGCAAACTGCTGACGTCAGGCTT
CTGGGGCTGGCCCGCCACTTCAACTACACCGGGGACCTCATGGCTCTCTAGCCCTACTGCGCGGCTGCGGATAGGCCACATCTCTCTTAC
TTCTACATCTGCTACATGACCACTCTGCTGGTCCACCGCTGTGTCGCTGAGCAGCAGCTGCAGCAGCAAGTACGGGAAGGACTGGAAGCGCT
ACACAGACGAGTTCCTTACCCAGTATTCTCTGGAGTTTGTAGTAAAGAGGATGAACCAAGTATTAGTCTCCACCGTTGCTTTAAGAAGG
GGTCTTTAGACCTATGCAAAACAAAAACATGCAAAACAAATGCAAGAAATAACCCAGTTCGCTTTACCAACATTTTACTTGAGTACATCT
CATGTGTTTAAATAAACAGATCCAGAGGACCTGTGGTCAAACATACATGACAGAAAGCCCTTCTTCTACTCTTCTTAAATACATTCGTGATG
TATTAAGAATACGTTTATCAAGAAACCCGTTGAGAAATCTCAGTTGGTACCTGATGGGACGACTGTCAGAAACCCGTTGGACCAAGCTCATGA
AGCAAAAGCAACTGATCTGAATTTCTTTGGATCCACTCCGTTTCTTCCGTCAGAGGACGAGCAAGCAGCCAGCTCTGCTGATGTTGGA
CATTTGACATCTTTGCTGTTTGTGAGCAGTTGGAACTTTTAGATTTACTGTAGAATTTGTAACAAAATCTTCAAATGTTTGTATAATTA
ACCAAAAGCAGACCCCTGAAAGCATAAAAAGCGCTTTTATACACTTTATGATGTGCTTTTTTTTTCTGTTCAATGCAATA

図1 Medaka genome database における *DHCR7* 遺伝子候補

黒に白抜き文字が、genome database によって推定されているメダカ *DHCR7* の exon 候補で、8カ所ある。GT は intron の始まり、AG は intron の終わりを示す。TAG は終始コドンを示す。ここで示す塩基配列は、メダカ第3染色体の、34524922 から 34533555 までの 8633bp を示す。

(a) Medaka genome database による推定 exon 2 候補の塩基配列:

GTTTGTCTATCTGCGGTTACCATGGAGACCACTAGGAGGGCGCGCGCAGCAGACGTCCAATGGGAAAGCCTCT
GAGCACGCGCAGCAGTCAGCACAGTGGGGCAGAGCATG (109bp)

(b) 本研究で Medaka genome database 推定 exon 2 候補中に見出した exon 1 候補の塩基配列 [(a) の下線部]:

ATGGAGACCACTAGGAGGGCGCGCGCAGCAGACGTCCAATGGGAAAGCCTCTGAGCACGCGCAGCAGTCAGC
ACAGTGGGGCAGAGCATG (89bp)

(c) 本研究で見出した exon 1 候補からの推定アミノ酸配列:

METTRRRRAQQ TSNGKASEHA QQSAQWGRA (29residues)

(d) メダカ, ヒト, ゼブラフィッシュ間での exon 1 アミノ酸配列の比較

MAA KSQLNIPKAK SLDGVTNDRT ASQGQWGRAW (Hs)
MMASDR VRKRHKGSAN GAQTVEKEPS KEPAQWGRAW (Dr)
METTRRRRAQQ TSNGKASEHA QQSAQWGRAW (Ol)

図2 メダカ DHCR7 exon 1 の推定

(a) 及び (b): Genome database での exon 1 及び exon 2 の塩基配列から, 実際の exon 1 の始まりは, genome database exon 2 の途中の ATG [(b) の四角枠] を開始コドンとして始まっていると仮定する。(c) 及び (d): すると, ヒトとゼブラフィッシュの exon 1 に共通に見られる C 末端の QWGRAW (黒に白抜き文字) というアミノ酸配列がメダカの場合でも見られるようになる。ただしメダカでは exon 1 の C 末端の W のコドンが次の exon とで分断されているので, この場合は W を exon 1 に含めた。(d) で, Hs はヒト, Dr はゼブラフィッシュ, Ol はメダカを示す。

(a) ヒト *Homo sapiens* の exon1 と exon2 の間での遺伝暗号 TGG の分割

```

                                     -----GCCC
ATG GCT GCA AAA TCG CAA CCC AAC ATT CCC AAA GCC AAG AGT CTA GAT
  M   A   A   K   L   Q   P   N   I   P   K   A   K   S   L   D

GGC GTC ACC AAT GAC AGA ACC GCA TCT CAA GGG CAG TGG GGC CGT GCC
  G   V   T   N   D   R   T   A   S   Q   G   Q   W   G   R   A
    
```

TG
 W(1st&2nd letters)

GT(635bp deleted from intron)AG (intron 639bp)

```

      G           GAG GTG GAC TGG TTT TCA -----
W(3rd letter) E   V   D   W   F   S -----
    
```

(b) ゼブラフィッシュ *Danio rerio* の exon1 と exon2 の間での遺伝暗号 TGG の分割

```

                                     -----AATC
ATG ATG GCA TCT GAC AGG GTG AGA AAG CGG CAT AAG GGC AGT GCC AAT
  M   M   A   S   D   R   V   R   K   R   H   K   G   S   A   N

GGT GCC CAG ACA GTG GAG AAG GAG CCA TCA AAG GAG CCG GCG CAG TGG
  G   A   Q   T   V   E   K   E   P   S   K   E   P   A   Q   W
    
```

GGA AGA GCA TG
 G R A W(1st&2nd letters)

GT(169bp deleted from intron)AG (intron 173bp)

```

      G           GAG GTG GAT TGG TTC TCT -----
W(3rd letter) E   V   D   W   F   S -----
    
```

図3 ヒトやゼブラフィッシュでもメダカと同様に exon1 の N 末端アミノ酸残基トリプトファン W の遺伝暗号は exon2 との間で分割されている

塩基配列のうち、遺伝暗号は3文字ずつ示し、相当するアミノ酸残基は遺伝暗号の下に1文字記号で示している。各 intron 内の配列は一部数字で、各 exon の前後の配列は一部・で示している。GT, AG は、GU-AG 則による intron の推定であることを示している。

表1 メダカ *DHCR7* 各 exon の塩基配列と推定アミノ酸配列

前の exon との境目にあるコドンで指定されているアミノ酸残基は () で閉じて、前の exon の最後に表示している。* は終止コドン。

	塩基配列	推定アミノ酸配列
exon1	ATGGAGACCACTAGGAGGCGCGCAGCAGACGTCCAATGGGAAAGCCTCTGA GCACGCGCAGCAGTCAGCACAGTGGGGCAGAGCATG (89 bp)	METTRRRRAQQTSNGKA SEHAQQSAQWGRA (W) (30)
intron	1620 bp	
exon2	GGAGGTGGACTGGTTCTCTCATCAGCGTGGTGGCTCCTCATCTTTGCTC CCTTCATCGTCTTTTACTTCGTGATGGCCTGTGACCAGTTCAGTGTCCGTC AGCGAGCCTTTGCTGCAGCTTTATTACGGGGAGCTGACGCTTCTCTCCATCTG GGCCCGGGCCCCCTCTTCTCTGGTGGCGGCTAAGATATACGCCGTTTGGG TGGGCTTCCAG (223 bp)	EVDWFSLISVVALLIF APFIVFYFVMACDQFQ CSVSEPLLQLLYGDVT LLSIWARAPSFWSAA KIYAVWVGFQ (74)
intron	77 bp	
exon3	GTCTTCCTCTACCTGTGTGTCCCTGATGTCACCCACAAGTTTATCCAGGGTA TGTTGGGGGGGTGCAGGATGGAGCACGAACCTCTGCAG (91 bp)	VFLYLCVPDVTCHKFIP GYVGGVQDGARTPA (G) (31)
intron	113 bp	
exon4	GTCTCATTAACAAGTATGAGATCAACGGGCTGCAGTGTGGCTGATCACGCAT GCCCTGTGGTTCGCCAACGCCGGTACTTCCACTGGTTTTCTCCCACCATCAT CTTCGATAACTGGATCCCCTGATGTGGTGCGCCAATTTGCTCGGCTACGCCG TGTCCACGTTTGCCTTCGTCAAGGCCTACATCTTCCCCACCAACCCCGAGGAC TG (214 bp)	LINKYEINGLQCWLIT HALWFANARYFHWFSF TIIFDNWIPLMWCANL LGYAVSTFAFKAYIF PTNPED (C) (71)
intron	79 bp	
exon5	CAAGTTCACAGGAAATGTCTTCTACAACACTACATGATGGGCATCGAGTTTAAC CGCGCATCGGCAAGTGGTTCGACTTCAAGCTTTTCTTCAACGGGCGACCCGGT ATTGTGGCCTGGACTCTAATCAATCTGTCTACATGTCCAAGCAGCAAGAGTT GTATGGAACCGTCACCAACTCCATGATTCTGGTCAACATCCTGCAG (205 bp)	KFTGNVFNYYMMGIEF NPRIGKWFDFKLFNG RPGIVAWTLINLSYMS KQELYGTVTNSMILV NILQ (68)
intron	1196 bp	
exon6	GGGATTTACGTGCTGGATTTCTTCTGGAATGAGGCGTGGTACCTGAAGACCAT CGACATCTGCCACGACCCTTCGGATGGTATCTGGGATGGGGGACTGCGTGT GGCTGCCATACCTGTACACGCTGCAG (132 bp)	GIYVLDFWNEAWYLK TIDICHDFGWYLGWG DCVWLPYLYTLQ (44)
intron	198 bp	
exon7	GGCCTGTACCTGGTTTACCACCCCGTCCAGCTCTCCGACGCTCCGCCCTCGC AGTCTGTGCTCGGCCTCTGGGATATTACATATTCCGCTCCACCAACCACC AGAAGGATCTCTCCGTCGCCACGATGGCTCTGCACAATTTGGGGCCGCAAG CCGTCTTTCATCGAATGCTCGTACCGCTCCGCCGATGGCGGCTTGCACCGAAG CAAACCTGCTGACGTGAGGCTTCTGGGGCGTGGCCCGCCACTTCAACTACCCG GGGACCTCATGGGCTCTTAGCCTACTGCGCCGCTGCGGATTAGGCCACATA CTTCCTTACTTCTACATCGTGTACATGACCATCTGTCTGGTCCACCGCTGTGT GCGTGACGAGCACCGCTGCAGCAGCAAGTACGGGAAGACTGGAAGCGCTACA CAGACGCAGTTCCTTACCGACTGATTCTGGAGTGTTTTAG (465 bp)	GLYLVYHPVQLSDASA LAVLLGLLGYIFRS TNHQKDLFRRHDGSC IWGRKPSFIECSYRSA DGGLHRSKLLTSGFWG VARHFNYTGDLMGSLA YCAACGLGHILPYFYI VYMTILLVHRCVRDEH RCSSKYGKDWKRYTDA VPYRLIPGVF* (154)

O1: METT**RRRA**QQT**SN**G**KAS**E**HA**Q**Q**S**AA**Q**W**G**RA**W**E**V**D**W**F**S**L**I**S**V**V**A**L**L**I**F**AP**F**I**V**F**Y**F**V**MA**C**D**Q**E**F**Q**
 Dr: MMASDRVRKRHKGS**AN**G**A**Q**T**V**E**K**E**P**S**K**E**P**AA**Q**W**G**RA**W**E**V**D**W**F**S**L**S**G**V**I**L**L**L**C**F**AP**F**L**V**F**F**F**I**MA**C**D**Q**Y**Q
 Hs: MAAKLQPNIPK**AK**S**L**D**GV**T**ND**R**TAS**Q**Q**W**G**R**A**W**E**V**D**W**F**S**L**A**S**V**I**F**L**L**L**F**AP**F**I**V**Y**F**I**MA**C**D**Q**Y**S**

O1: **CS**V**SE**P**LL**Q**LY**Y**GD**V**T**L**S**I**WA**R**AP**S**F**S**W**S**AA**K**I**Y**AV**W**V**G**F**Q**V**E**LY**L**CV**P**D**V**T**H**K**F**I**P**GY**V**GG**V**Q**D**G**
 Dr: **CS**I**SH**P**LL**D**LY**N**G**D**AT**L**FT**I**WN**R**AP**S**F**T**WA**A**AK**I**Y**A**I**W**V**T**F**Q**V**V**LY**M**CV**P**D**F**L**H**K**I**L**P**GY**V**GG**V**Q**D**G**
 Hs: **C**A**L**T**G**P**V**V**D**I**V**T**G**H**A**R**L**S**D**I**WA**K**T**P**P**I**T**R**K**A**A**Q**LY**T**L**W**V**T**F**Q**V**L**LY**T**S**L**P**D**F**C**H**K**F**L**P**GY**V**GG**I**Q**E**G

O1: **AR**T**P**A**G**L**I**N**K**Y**E**I**N**G**L**Q**C**W**L**I**T**H**A**L**W**F**A**N**A**R**Y**F**H**W**F**S**P**T**I**I**F**D**N**W**I**P**L**M**WC**A**N**L**L**G**Y**A**V**S**T**F**A**F**V**K**A**
 Dr: **AR**T**P**A**G**L**I**N**K**Y**E**V**N**G**L**Q**C**W**L**I**T**H**V**L**W**V**L**N**A**Q**H**F**H**W**F**S**P**T**I**I**I**D**N**W**I**P**L**L**W**C**T**N**I**L**G**Y**A**V**S**T**F**A**E**I**K**A
 Hs: **A**V**T**P**A**G**V**V**N**K**Y**Q**I**N**G**L**Q**A**W**L**L**T**H**L**L**W**F**A**N**A**H**L**L**S**W**F**S**P**T**I**I**F**D**N**W**I**P**L**L**W**C**A**N**I**L**G**Y**A**V**S**T**F**A**M**V**K**G**

O1: **Y**I**F**P**T**N**P**E**D**C**K**F**T**G**N**V**F**Y**N**Y**M**M**G**I**E**F**N**P**R**I**G**K**W**F**D**F**K**L**F**F**N**G**R**P**G**I**V**A**W**T**L**I**N**S**Y**M**S**K**Q**Q**E**L**Y**G**T**V
 Dr: **Y**L**F**P**T**N**P**E**D**C**K**F**T**G**N**M**F**Y**N**Y**M**M**G**I**E**F**N**P**R**I**G**K**W**F**D**F**K**L**F**F**N**G**R**P**G**I**V**A**W**T**L**I**N**S**Y**A**A**K**Q**Q**E**L**Y**G**Y**V
 Hs: **Y**F**F**P**T**S**A**R**D**C**K**F**T**G**N**E**F**Y**N**Y**M**M**G**I**E**F**N**P**R**I**G**K**W**F**D**F**K**L**F**F**N**G**R**P**G**I**V**A**W**T**L**I**N**S**F**A**A**K**Q**R**E**L**H**S**H**V

O1: **T**N**S**M**I**L**V**N**I**L**Q**C**I**Y**V**L**D**F**F**W**N**E**A**W**Y**L**K**T**I**D**I**C**H**D**H**F**G**W**Y**L**G**W**G**D**C**V**W**L**P**Y**L**T**L**Q**G**L**Y**L**V**Y**H**P**V**Q**L**S
 Dr: **T**N**S**M**I**L**V**N**V**L**Q**A**V**Y**V**V**D**F**F**W**N**E**A**W**Y**L**K**T**I**D**I**C**H**D**H**F**G**W**Y**L**G**W**G**D**C**V**W**L**P**F**L**Y**T**L**Q**G**L**Y**L**V**Y**N**F**I**Q**L**S**
 Hs: **T**N**A**M**V**L**V**N**V**L**Q**A**I**Y**V**I**D**F**F**W**N**E**T**W**Y**L**K**T**I**D**I**C**H**D**H**F**G**W**Y**L**G**W**G**D**C**V**W**L**P**Y**L**T**L**Q**G**L**Y**L**V**Y**H**P**V**Q**L**S

O1: D**A**S**AL**A**V**L**L**L**G**L**L**G**Y**Y**I**F**R**S**T**N**H**Q**K**D**L**F**R**R**H**D**G**S**C**T**I**W**G**R**K**P**S**F**I**E**C**S**Y**R**S**A**D**G**L**H**R**S**K**L**L**T**S**G**F**W
 Dr: **T**P**H**A**A**G**V**L**I**L**G**L**V**G**Y**Y**I**F**R**V**T**N**H**Q**K**D**L**F**R**R**T**E**G**N**C**S**I**W**G**K**K**P**T**F**I**E**C**S**Y**Q**S**A**D**G**A**I**H**K**S**K**L**M**T**S**G**F**W**
 Hs: **T**P**H**A**V**G**V**L**L**L**G**L**V**G**Y**Y**I**F**R**V**A**N**H**Q**K**D**L**F**R**R**T**D**G**R**C**L**I**W**G**R**K**P**K**V**I**E**C**S**Y**T**S**A**D**G**Q**R**H**H**S**K**L**I**V**S**G**F**W**

O1: **G**V**A**R**H**F**N**Y**T**G**D**L**M**G**S**L**A**Y**C**A**C**G**L**G**H**I**L**P**Y**F**Y**I**V**Y**M**T**I**L**L**V**H**R**C**V**R**D**E**H**R**C**S**S**K**Y**G**K**D**W**K**R**Y**T**D**A**V**P
 Dr: **G**V**A**R**H**M**N**Y**T**G**D**L**M**G**S**L**A**Y**C**L**A**C**G**G**N**H**L**L**P**Y**F**Y**I**I**Y**M**T**I**L**L**V**H**R**C**I**R**D**E**H**R**C**S**N**K**Y**G**K**D**W**E**R**Y**T**A**A**V**S**
 Hs: **G**V**A**R**H**F**N**Y**V**C**D**L**M**G**S**L**A**Y**C**L**A**C**G**G**G**H**L**L**P**Y**F**Y**I**I**Y**M**A**I**L**L**T**H**R**C**L**R**D**E**H**R**C**A**S**K**Y**G**R**D**W**E**R**Y**T**A**A**V**P**

O1: **Y**R**L**L**I**P**G**V**F***
 Dr: **Y**R**L**L**P**N**I**F*
 Hs: **Y**R**L**L**P**G**I**F*

図4 メダカ *DHCR7* 推定アミノ酸配列のヒトやゼブラフィッシュとの比較

O1, Dr, Hsは略号で図2に同じ。黒に白抜き文字は、3種の生物のうち、2種以上で同じアミノ酸残基であることを示す。灰色に黒文字は、類似の性質を示すアミノ酸残基であることを示す。四角枠は、*DHCR7* 遺伝子のコンセンサス配列であるS-1配列を示す。各アミノ酸配列は、National Center for Biotechnology InformationのAccession numberが、HsがNM_001360、DrがNM_201330のデータから引用している。

