

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 *Squalene epoxidase* の遺伝子構造～

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

西野 秀昭

田中 哲

Hideaki NISHINO

Satoshi TANAKA

福岡教育大学・理科教育講座 (生物)

(平成24年10月1日受理)

Abstract

The gene structure of Medaka *Oryzias latipes Squalene epoxidase* 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, the all possible exon structure, 1st exon to 11th exon consisting of 1716 bp open reading frame and the estimated 572 amino acid residues, of *Squalene epoxidase* of Medaka *Oryzias latipes* were extracted as the strong candidates in the 16th chromosome data base. The deleted parts of the 1st and 2nd exons in *Danio rerio Squalene epoxidase* were also found in comparison to the enzyme of *Homo sapiens* and *Oryzias latipes*.

要 約

ステロール生合成遺伝子が、生物進化の過程で如何にして創生されてきたか探究するため、ステロール生合成経路の調節段階でもある酵素をコードする *Squalene epoxidase* 遺伝子の構造を複数の生物で比較する目的で、メダカ *Oryzias latipes* の *Squalene epoxidase* 遺伝子構造の解明に取り組んだ。ヒト *Homo sapiens* やゼブラフィッシュ *Danio rerio* の *Squalene epoxidase* の cDNA 構造でナショナルバイオリソースプロジェクト・メダカゲノムデータベースを対象に BLAST 検索を行った。その結果、第16番染色体を BLAST 検索の対象とすると、メダカ *Oryzias latipes* の *Squalene epoxidase* の第1から第11exon の候補全てを見出すことができた。読み枠の塩基対の数は終止コドンを除き 1716、推定アミノ酸配列の残基数は 572 個であった。11 個という exon 数は *Homo sapiens*, *Danio rerio* と同じであった。一方、第2exon で、*Homo sapiens* や *Oryzias latipes* と比較して *Danio rerio* では欠失していると考えられる塩基配列が見いだされた。推定アミノ酸配列も第1exon と第2exon 以外は高度に類似していた。これらのことから、メダカゲノムデータベースから得られた塩基配列は、メダカ *Squalene epoxidase* 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ることから、生物が前核細胞から真核細胞へと進化する際に、何らかの過程によって新しい遺伝子として誕生してきたものと考えられる¹⁾。ステロール構造が存在することによって、細胞膜の構造は柔軟性を増すとともに、細胞膜が柔軟になることによって、細胞の大型化や貪食能の獲得へとつながったことが推定されている²⁾。

ステロール生合成遺伝子の一つ、*Squalene epoxidase* は、コレステロール生合成過程の Squalene 閉環段階を触媒する酵素をコードしている³⁾。動物では、*HMG-CoA reductase* とともに、コレステロール生合成経路の律速段階の一つとして考えられている³⁾。

メダカ (*Oryzias latipes*) は、我が国で独自に開発された脊椎動物のモデル生物として、生物学、医学、環境科学、水産学などで広く利用され、英語名も Medaka と表記されている⁴⁵⁾。さらに、比較ゲノム生物学の進展によって、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期待されている⁵⁾。したがって、例えば、メダカの遺伝的な変異体を解析することによって、ヒト遺伝病の原因遺伝子にたどりつく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ており⁵⁾、モデル生物でも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なっている。またメダカは、日本国の自生種であることから、野生集団の遺伝子解析を行うことによって、我が国の自然環境の成り立ちや、その保全に新たな視点を提供する可能性も考えられる。

本研究の目的は、メダカの *Squalene epoxidase* 遺伝子構造を明らかにすることである。それは、ステロール生合成遺伝子が生物の進化の過程でどのようにして誕生してきたのかを考察する上で意義ある研究になるものと考えられるからである⁴⁾。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)⁶⁾ 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 *Squalene epoxidase* の情報は未だ見られないが、我が国のナショナルバイオリソースプロジェクト (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⁵⁾。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⁵⁾。しかし、少なくとも *Squalene epoxidase* の遺伝子構造は、その cDNA の塩基配列とともに未だ同定されていない。

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ゲノムデータや mRNA 配列データが NCBI に登録されている。ヒト *Squalene epoxidase* 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 *Squalene epoxidase* 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 *Squalene epoxidase* でも同様に類似していると考えられることができる。

本研究では、NCBI からヒト *Squalene epoxidase* を含むゲノムと cDNA 及び、ゼブラフィッシュ *Squalene epoxidase* を含むゲノムと cDNA の各情報を利用して、NBRP メダカゲノムプロジェクトのデータとの間での類似性の検索、すなわち Basic Local Alignment Search Tool (BLAST) によって検索し、メダカ *Squalene epoxidase* 遺伝子構造を見出し、必要に応じて塩基配列のクローニングへ結びつけることをも視野に入れながら、コレステロール生合成遺伝子の創生メカニズムの解析へ供することを目的とした。

2. 研究方法

NCBI 法: 既に全塩基配列が決定されデータベースに登録されている *Squalene epoxidase* の塩基配列のうち、ヒトと硬骨魚類、ここではゼブラフィッシュで同じアミノ酸配列を探し出す。ヒトと硬骨魚類であるゼブラフィッシュで共通なアミノ酸配列は、硬骨魚類であるメダカでも共通である事が期待されるからである。そこでまず、ヒト及びゼブラフィッシュの各 cDNA (mRNA) 塩基配列を、遺伝情報データベース GenBank から探し出す。米国国立衛生研究所 (NIH) が運営している Web サイト、National Center for Biotechnology Information (NCBI) の HP を開く (<http://www.ncbi.nlm.nih.gov/>)。そのサイトにおいて、「Nucleotide」を選択し、*Homo sapiens Squalene epoxidase* をキーワードとして検索を実行する。例えば、*Homo sapiens Squalene epoxidase* の cDNA 情報は、accession number が NM_003129.3

GI:62865634 [Homo sapiens Squalene epoxidase (SQLE)] というデータに登録されている。次に cDNA (mRNA) とゲノム DNA の塩基配列を比較して, exon 構造を塩基配列レベルで見出す。ヒトゲノムにおける *Squalene epoxidase* の塩基配列は, accession number が AF098865 (NC_000008) の遺伝子情報を用いる。Danio rerio のゲノム遺伝子情報は, NC_007127.4 のものを用いた。真核生物のゲノムには, 最終の RNA 産物に残る exon 構造と, 最初の転写産物より除かれる intron 構造が交互に配置された, 分断された遺伝子構造である^{14,7)}。exon 構造は, 完成した成熟 mRNA に残される塩基配列である。このように遺伝子は, mRNA の 5' 末端にあたる exon にはじまり, 3' 末端にあたる exon で終わる。intron は介在配列であり, 一次転写産物がプロセッシングを受けて完成した mRNA になる際に切り除かれる (RNA スプライシング⁷⁾)。異なる生物間で相同遺伝子を比較した場合, 通常 intron の位置は保存されているが, それぞれの intron の長さはかなり変化している。しかし, exon 配列はよく似たままで保存されている。exon の配列が保存されていることを利用して, 異なった種からよく似た遺伝子や塩基配列を見出すことができる。RNA スプライシングは, intron のスプライス部位で行われる。スプライス部位は, exon-intron 境界の intron 側にある配列で, intron の 5' 末端に位置する 5' スプライス部位のコンセンサス配列は GU から始まり, intron の 3' 末端に位置するスプライス部位のコンセンサス配列は AG で終わる^{14,7)}。このように, intron 構造は原則として, GU で始まり, AG で終わるので, これを GU-AG 則と呼ぶ事とする¹⁾。従って exon 構造を見つけるには, intron 構造が GU と AG で挟まれていることを目印としてスプライシングされることから, cDNA (mRNA) の塩基配列とゲノムにおける遺伝子の塩基配列を比較することで exon-intron 構造を推定できる。そして, ヒトの *Squalene epoxidase* と, ゼブラフィッシュの *Squalene epoxidase* を比較し, 類似性が見出されれば, ゼブラフィッシュと同じ硬骨魚類であるメダカの *Squalene epoxidase* もヒト *Squalene epoxidase* やゼブラフィッシュ *Squalene epoxidase* のアミノ酸配列に類似していることが期待できる。その類似した配列を基に, メダカ *Squalene epoxidase* の遺伝子構造を見出す事が可能と考えられる。その方法が次に示す検索ツール BLAST による方法である。

BLAST 法 : メダカ *Oryzias latipes* は我が国で独自に開発された脊椎動物のモデル生物として, 生

物学, 医学, 環境科学, 水産学などで広く利用されている。さらに比較ゲノム生物学の進展によって, メダカとヒトのゲノムには多くの共通性があることが明らかになり, メダカは遺伝子機能解析の重要なバイオリソース (生物資源) となっている。「ナショナルバイオリソースプロジェクト (NBRP)」は, ライフサイエンスの研究に広く用いられる実験材料としてのバイオリソース (実験動物, 細胞, DNA などの遺伝子材料で, ここでは「生物遺伝資源」と同義語として扱う)のうち, 国が特に重要と認めたものについて, 体系的な収集, 保存, 提供体制を整備することを目的とした国家プロジェクトである⁵⁾。またバイオリソースに関する情報も同様に収集し, 使いやすいデータベースとして利用者に提供することによって, リソースの有効利用を図ることも目的としている。

まず最初に, NBRP のホームページサイト (<http://www.nbrp.jp/index.jsp>) を開く。左に多数のバイオリソース名があるので, その中の「メダカ」をクリックする。「NBRP メダカ Genome」をクリックする。「BLAST」をクリックする。「BLAST」とは, 先にも述べたように, Basic Local Alignment Search Tool の略で, 類似した塩基配列やアミノ酸配列をデータベース等から見つけることができる web 上の検索ツールである。GenBank に登録されている遺伝子情報は, 「DNA sequence」から検索することができる。NBRP で公開されているメダカゲノムデータは, 平成 14 年度文部科学省ナショナルバイオリソースプロジェクト「ゲノム解析等」によって実施されたニホンメダカ *Oryzias latipes* のゲノム解析によって得られたもので, 解析にはメダカ近交系 Hd-rR が用いられている。この系統は, 放射線医学総合研究所の石川裕二博士から提供されたものである。また, その DNA は東京大学の武田洋幸博士から提供され, ゲノム解析は国立遺伝学研究所および理化学研究所によって, ホールゲノムショットガン法によって行われた。約 100 万リード, 726Mbp の情報が含まれている。メダカのゲノムサイズは約 800Mb と推定されているので, 全ゲノムの 0.9 × 分に相当する⁵⁾。

BLAST のプログラムは, 以下の 3 つがある。

- ① blastn : DNA query (検索したい塩基配列) を DNA database に対して行う
- ② blastb : タンパク質に翻訳された DNA query を, 同様にタンパク質に翻訳された DNA database に対して行う
- ③ blastp : protein query を タンパク質に翻訳さ

れた DNA database に対して行う

このうち、使用したプログラムは、①の blastn と③の blastp である。blastn と blastp の使い分けは、cDNA の Open Reading Frame (ORF, 翻訳可能な読み枠) の塩基配列や mRNA 配列を BLAST する場合は blastn を、アミノ酸配列を BLAST するときは blastp を用いた。blastn では、ある塩基配列とよく似た塩基配列を、似ている順に複数見つけることができるもので、blastp は、あるアミノ酸配列とよく似たアミノ酸配列を、似ている順に見つける事ができる。②の blastb は、ある塩基配列をタンパク質に翻訳して、よく似たアミノ酸配列を検索することができるが、paste した最初の塩基から翻訳を開始し、順番に翻訳してしまうので、求めるアミノ酸配列と異なるものを検索してしまう恐れがあるため、ここでは使用しなかった。

まず、blastn では、「COPY&PASTE」の枠に、ヒト *Homo sapiens Squalene epoxidase* mRNA の塩基配列を DNA に置き換えたもの、すなわち cDNA の ORF を paste して、「Search」をクリックする。タンパク質に翻訳される領域では、コドンの塩基が一つ異なってもアミノ酸残基が同じであれば、タンパク質として同じ働きをすることから、塩基配列よりもホモロジーを見つけやすいアミノ酸配列で検索、すなわち blastp も行った。ヒト *Homo sapiens Squalene epoxidase* mRNA をアミノ酸に翻訳したもので検索を行った。ゼブラフィッシュ *Danio rerio* についても、同様に *Squalene epoxidase* mRNA をアミノ酸に翻訳したもので BLAST 検索した。ヒトとゼブラフィッシュの検索結果で、同じデータがヒットしていれば、それはメダカ *Oryzias latipes* の *Squalene epoxidase* である可能性が高いと考えられる。

次に、exon ごとにも BLAST 検索を行った。すなわち、ヒト *Homo sapiens Squalene epoxidase* 各 exon、ゼブラフィッシュ *Danio rerio Squalene epoxidase* 各 exon をメダカゲノム情報で BLAST 検索した。その際、各 exon の塩基配列で検索すると共に、各 exon の塩基配列をアミノ酸に変換したもので検索を行った。得られたデータから、メダカ *Oryzias latipes Squalene epoxidase* の exon である可能性が高いものを見出していった。intron は GU-AG 則によると、GU と AG で挟まれている。これを目印にして未成熟の mRNA はスプライシングされ成熟した mRNA となる。よってデータでヒットしているアミノ酸配列に相当する塩基配列の前後で GU-AG 則に一致する配列を

見つけた。

ヒトやゼブラフィッシュの遺伝子構造との比較から BLAST 検索によっても見つけられていない exon が存在する可能性が高い場合には、メダカのどの染色体上にある可能性が高いか検討した。

3. 結果と考察

メダカ *Squalene epoxidase* の cDNA や遺伝子構造情報の検索：まず、メダカ *Squalene epoxidase* の研究がどのくらいなされているのかを調べるために、NIH の National Center for Biotechnology Information (NCBI) にて情報の検索を行い、DNA 配列、mRNA 配列情報を探した⁴⁾。検索ページの枠部分に '*Oryzias latipes Squalene epoxidase*' と入力し、search を選択し、メダカ *Oryzias latipes Squalene epoxidase* に関する情報の検索結果を表示させた。この画面のデータベースを示す 'Nucleotide' には遺伝子の DNA や mRNA の情報が登録されていて、'Protein' にはアミノ酸配列情報が登録されている。しかし、*Oryzias latipes Squalene epoxidase* を用いた検索では、この二つの項目でヒットした情報はなかった。*Squalene epoxidase* は、*Squalene monooxygenase* (または *monooxygenase*) という別名があるので、'*Oryzias latipes Squalene monooxygenase*' または '*Oryzias latipes Squalene monooxygenase*' で検索したが、同じようにヒットする情報はなかった。つまり、メダカ *Oryzias latipes Squalene epoxidase* に関する研究の情報がまだ登録されていないと考えられた。このように、メダカ *Oryzias latipes Squalene epoxidase* の遺伝子情報は NCBI のデータベース中には見出すことができなかつたので、少なくとも本研究でメダカ *Oryzias latipes Squalene epoxidase* の遺伝子構造を調べることは価値があるものと判断した。

ヒト及びゼブラフィッシュ *Squalene epoxidase* の cDNA や遺伝子構造情報の検索：本研究ではまず、メダカ *Oryzias latipes Squalene epoxidase* の遺伝子情報を、他の脊椎動物のデータと比較対照しながら検索するため、ヒト *Homo sapiens Squalene epoxidase* の遺伝子情報と、メダカと同じ淡水系の硬骨魚類であるゼブラフィッシュ *Danio rerio* の *Squalene epoxidase* の遺伝子情報を NCBI のデータベースの中から検索した。すなわち、ヒト *Homo sapiens Squalene epoxidase* とゼブラフィッシュ *Danio rerio Squalene epoxidase*

の cDNA と遺伝子構造情報をもとに、メダカ *Oryzias latipes Squalene epoxidase* を BLASAT で検索していく。

ゼブラフィッシュ *Squalene epoxidase* は 'Danio rerio Squalene epoxidase' で、ヒト *Squalene epoxidase* は 'Homo sapiens Squalene epoxidase' で NCBI ホームページでの検索を行った。データ内容を見るには、選択して情報が載っているページを開いた。ヒト *Squalene epoxidase* の cDNA 情報は、accession number AF098865 を用いた。この mRNA 情報とゲノム DNA の塩基配列を比較して、exon 構造を求めた (図 1)。異なる生物間で相同遺伝子を比較した場合、通常 intron の位置は保存されているが、それぞれの intron の長さや塩基配列ははかかなり変化している¹⁾。しかし、exon の塩基配列はよく似たままで保存されている¹⁾。exon の配列が生物進化の上で保存されていることを利用して、異なった種からよく似た遺伝子を分離することができることが期待される。GU-AG 則を目印としてスプライシングされることから、mRNA と比較して見つけることができる。ゼブラフィッシュについても同様に exon を見つけた (図 2)。

遺伝子構造としてヒト *Squalene epoxidase* と、ゼブラフィッシュ *Squalene epoxidase* を比較すると、exon の数は共に、11 個であった (図 1, 図 2)。各 exon でのアミノ酸の数は exon1, exon2, exon3 ではいずれもヒト *Squalene epoxidase* のほうが多かったが、exon4 ~ exon11 のアミノ酸の数は同じであった (図 1, 図 2)。Intron の入り方に違いが見られたが、exon の塩基数はほぼ同じであり、exon の塩基配列も比較的、中立進化的に保存されていることが分かる。また、全アミノ酸配列も同様に比較すると、ヒト *Homo sapiens* とゼブラフィッシュ *Danio rerio* で高度に類似していた。

このようにアミノ酸配列も、ヒト *Squalene epoxidase* とゼブラフィッシュ *Squalene epoxidase* では相互に類似しているため、メダカ *Squalene epoxidase* も、ヒト *Squalene epoxidase* やゼブラフィッシュ *Squalene epoxidase* に似ていることが期待できる。従って、ゼブラフィッシュ *Squalene epoxidase*、ヒト *Squalene epoxidase* の塩基配列やアミノ酸配列とよく似た配列をメダカゲノムから見いだすことによって、メダカ *Squalene epoxidase* の遺伝子構造を見つめることができると思われる。その方法が「BLAST」検索である。

National Center for Biotechnology Information

(NCBI) では、メダカ *Squalene epoxidase* の cDNA (mRNA) 塩基配列やタンパク質の登録データを見出すことができなかった。そこでメダカの *Squalene epoxidase* 遺伝子情報を見つけるために、National Bio Resource Project (NBRP) のウェブサイトを利用し検索を行うことにした。以下に、行った手順を示す。

NBRP のホームページを開き、メダカを選択した (下図, NBRP 情報公開サイトより引用)。



次にデータベースページを選択し、NBRP Medaka ページへ移動した (下図, NBRP 情報公開サイトより引用)。



「Blast 検索」を選択し、BLAST 検索のページへと移動した。そのページの「種類」を選択し、「枠」には BLAST 検索にかけたいデータを入力した。ここでも検索プログラムは、以下の 3 つ、Blastn, Blastb, Blastp, がある。本研究で使用した検索プログラムは、Blastp (protein query vs. DNA database[translated into protein]) であり、アミノ酸配列をペーストして BLAST 検索を行った。

BLAST 検索にあたっては、例えば、ゼブラフィッシュ *Squalene epoxidase* の各エクソンがコードするアミノ酸配列を用いた BLAST 検索では、*Danio rerio* の exon 1 のアミノ酸配列、MWTFLGIASF TYIYKCDAL LSSATTELLV

AAVLCVSVGL IITCFSLRGQ TLKAPQFLQM PFQLLMTFSP TKELFSASST DSCSRSSKK を、ペーストした。BLAST 検索の結果、表 1 に示すように exon1 ~ exon3, exo8 ~ exon11 までは、BLAST 検索でヒットしたデータをもとにメダカゲノム DNA から *Squalene epoxidase* の exon を見つけることができた。しかし、exon4 ~ exon7 は BLAST 検索で有効なヒットが観察されず、存在を確認することができなかった。

BLAST 検索による結果から、メダカの第 16 番染色体上にヒットしたデータ、すなわちリードと関連する部分が多かったことから、メダカ *Squalene epoxidase* 遺伝子が、第 16 番染色体上に存在する可能性が高いものと考えた。イントロン部分は、RNA ではコンセンサス配列 GU と AG には含まれている。つまり、DNA では GT と AG には含まれていることになるので、その部分に注意しながらメダカの exon 4 ~ exon 7 を見出した (図 3, 表 2, 図 4)。

BLAST 検索では、exon ごとのアミノ酸配列による検索で、ヒト *Homo sapiens*, ゼブラフィッシュ *Danio rerio* の *Squalene epoxidase* に共通してヒットしたものから類似性の高い配列を見つけることで、メダカ *Oryzias latipes* の *Squalene epoxidase* 遺伝子情報を見つけることができる可能性が考えられること、メダカのはほすべてのゲノム塩基配列が決定されたデータベースが NBRP によって整備されていることから、本研究はスタートした。

ヒトとゼブラフィッシュで共通してヒットしたリードの中に、*Oryzias latipes* のゲノム情報と思われるものがなかった場合や、共通してヒットしなかった exon については、以前は *Oryzias latipes* と近縁な生物のゲノム情報を用いて検索するなどの BLAST 方法の必要性を考えていたが、ヒットした exon が存在する染色体の塩基配列情報があれば、本研究のように一部の exon しか BLAST 検索で見出せなくても、その塩基配列をきっかけにして求める遺伝子が存する可能性が高い染色体を見つけ、GU-AG 則や他の生物の遺伝子情報との比較から見出すことが可能となることを示せた。

本研究で、NBRP メダカゲノム情報データベースから推定したメダカ *Oryzias latipes Squalene epoxidase* の exon が実際に存在しているのか、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はしていない。今後この確認を行うことによって推定したメダカ *Oryzias latipes Squalene epoxidase* の遺伝子構造を確認しなければならない。

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) Sakakibara, J., Watanabe, R., Kanai, Y., & Ono, T., Molecular Cloning and Expression of Rat Squalene epoxidase, J. Biol. Chem., Vol. 270, No. 1, January 6, 17-20, 1995
- 4) 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
- 5) ナショナルバイオリソースプロジェクト (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
- 6) National Center for Biotechnology Information (NCBI) HP, <http://www.ncbi.nlm.nih.gov/>
- 7) RNA splicing : Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P.: MOLECULAR BIOLOGY OF THE CELL, Chapter 6, HOW CELLS READ THE GENOME: FROM DNA TO PROTEIN, 299-374, 2002

(a) ヒト *Homo sapiens Squalene epoxidase* 各 exon の塩基配列 (1722 塩基対)

exon1: atgtggacttttctgggcattgccactttcacctatttttataagaagttcggggacttcatcactttggccaacagggaggtcctgt
 tgtgctgctggtgttctctcgtcggcctggtgctctcctaccgctgtcgccaccgaaacgggggtctcctcgggcgcccagcagag
 cggctcccagttcgcctcttctcggatattctctcaggcctgccttcatgtgcttctctgggccaatccccctgaatcagaa
 aataaggagcagctcgaggccaggagg (291)

exon2: cgcagaaaaggaaccaatatttcagaaaacaagcttaataggaacagctgcctgtacatcaacatcttctcagaatgaccagaagtta
 tcatcgtgggagctggcgtgcttggctcgttggcagctgtgcttccagagatggaagaaggtgacagtcattgagagagactt
 aaaagagcctgacagaaatagttggagaattcctgcagccgggtggttatcatgttctcaaagaccttggcttggag (253)

exon3: atacagtggaaaggtccttgatgccaggttgtaaattggttacatgattcatgatcaggaagcaaatcagaggttcagattccttacc
 tctgtcagaaaaacaatcaagtgcagagtggaaagactttccatcacggaagattcatcatgagtctccggaaagcagctatggcagag
 cccaa (181)

exon4: tgcaaagtttatgaaggtgtgtgttacagttattagaggaagatgatgtgtgatgggagttcagtaacaggataaagagactgga
 gatatacaag (97)

exon5: gaactccatgctccactgactgttgttgcagatgggcttttctccaagttcaggaaaagcctggctccaataaagtttctgtatcat
 ctcatttgttggcttcttattgaag (114)

exon6: aatgcaccacagtttaagcaaatcatgctgaacttatttagtaaccgagtcagttctcatctaccagatttcatccagtgaaa
 ctcgagtacttgttgacattagaggagaaatgccaaaggaatttaagagaatacatggttgaaaaaattaccacaaaatacctg
 (172)

exon7: atcacctgaaagaaccattcttagaagccactgacaattctcatctgaggtccatgccagcaagcttcttctccttcatcagtgaa
 gaaacgag (96)

exon8: gtgttcttcttttgggagacgcatataatagggcatccacttactggtggaggaatgactgttgcttttaagatataaaaactatg
 gagaaaactgctaaaggtatccctgaccttattgatgatgcagctattttctgag (143)

exon9: gccaaaaaatcattttactgggcaagaaaaacatctcattcctttgtcgtgaatatccttgctcaggctctttatgaattattttctg
 ccacagatg (97)

exon10: attcctgcatcaactaagaaaagcctgttttctttatttcaacttgggtggcaatgtgttgcgggtcctgttgggctgctttctg
 t (88)

exon11: attgtctcctaaccctctagttttaaattggacacttcttctgctgttgaatctatgcctgtatttttgctttaagtacagaaacctg
 gattacaaaacctcgagccttctcagtagtgggtgctgtattgtacaaagcgtgttctgtaaatatttctctaatttactcagaaat
 gaagtatatggttcat (190)

(b) ヒト *Homo sapiens Squalene epoxidase* の各 exon におけるアミノ酸配列 (574 アミノ酸残基)

exon1: MWTFLGIATF TYFYKKFGDF ITLANREVLL CVLVFLSLGL VLSYRCRHRN GLLGRQSQS SQFALFSDIL SGLPFIGFFW
 AKSPPESENK EQLEARR (97)

exon2: RRK GTNISETSLI GTAACSTSS QN DPEVIIVGAG VLGSALAAVL SRDGRKVTVI ERDLKEPDRI VGEFLQPGGY RVLKDLGLG
 (84)

exon3: (D) TVEGLDAQ VVNGYMIHDQ ESKSEVQIPY PLSENNQVQS GRAFHHRFI MSLRKAAMAEP (60)

exon4: (N) AKFIEGVV LQLEEDDVV MGVMYKDKET GDIK (33)

exon5: ELHAPL TVVADGLFSK FRKSLVSNKV SVSSHVGFGL IK (38)

exon6: NAPQFKAN HAELILANPS PVLIIQISSS ETRVLVDIRG EMRNLREYM VEKIYPQIP (57)

exon7: (D) HLKEPFLEAT DNSHLRSMFA SFLPPSSVKK R (32)

exon8: (G) VLLLGDAY NMRHPLTGGG MTVAFKDIKL WRKLLKGIPD LYDDAAIFE (48)

exon9: AKKSFYWARD TSHSFVNNIL AQALYELFSA TD (32)

exon10: (D) PLHQLRKA CFLYFKLGGG CVAGPVGLLS (29)

exon11: (V) LSPNPLVLI GHFFAVAIYA VYFCFKSEPW ITKPRALLSS GAVLYKACSV IFPLIYSEMK YMVH (64)

図1 ヒト *Homo sapiens Squalene epoxidase* 遺伝子構造における exon 塩基配列とアミノ酸配列

() 内のアミノ酸一文字記号は、前の exon とで分断された遺伝暗号によるアミノ酸残基を示す。その場合、分断の後の exon でのアミノ酸残基としている。() 内の数字は、塩基対数またはアミノ酸残基数を表す。

(a) ゼブラフィッシュ *Danio rerio* *Squalene epoxidase* 各 exon の塩基配列 (1671 塩基対)

```

exon1:atgtggacctttcttggcatagcgagttttacatatactataagaaatgcgacgcgctgctgagctccgcgactacagagctgctgg
    tggccgcggtgctgctgctgctgctgctgctggggctcatcatcacctgtttctccctccgagcagactctcaaagcgccagtttttaca
    aatgccttttcaactcttgatgacattttccccaccaaaaaactcttctccgcgctccagcactgatagttgtagtgcctcttaag
    aag (267)
exon2:agagatgtaaagcgaagaggggcgacggacagtcagtgagtcgactgagtgatccgagggatccgagggatcatcatagtgaggagctggggtgctag
    gatcgccatggctgctgttctggcccggtggagcggcgggtcaccgctggaggggacatgaaagaaccagaccgaatagttgg
    agaactctgagccgggaggatttccgggtactcaaagaccctggcctggagg (229)
exon3:gtactgttgagggtcctgagctgctcatggtggaacggttatgtaatccatgacatggaagttaggacagaggtggagatcccataatcc
    ccagcaggagaactgcattcaggggacgtgctttccatcatgggcgcttcatcatggggtgcgctgcgagctgccctcgctgagcct
    aa (178)
exon4:tgtgaagtttgtagggaacagtgaccagcctggagaggaggaaggttgtgtgacaggaattcagtacaggagaaagacaccggg
    atgctcaag (97)
exon5:gaggttcgcgctcctctgactatagtcgctgatggtgcttttcccaagttcaggaagaatctggtttctgggaaagttcgggtgctg
    cgattttggtgatgcataatgaag (114)
exon6:gactctcctcagttcaaagcccaccatgcagagctagttctagctgatcccagtcagtcgctgactcaccagatctcctccaatgaaa
    ctgctgttctggtggacatcagaggagagatgccacgggacctcatgacgtacatgactgacaaaactaccacagctaccag
    (172)
exon7:agcacttaagagagccgttctgctggcactgcagaacgaccggttgagaacgatgcctgccagctttctgcccaccgcaccagtc
    aaaaacgaag (96)
exon8:gggttctgctgttaggagatgcgtataacatgcgtcatcctctgaccggcggaggatgagtgctgctgaatgacgtctccatctg
    gagagacctgctcaagaacattcctgacacctatgacaatacagcttctgctccag (143)
exon9:gcgaagaagaaattcactgggagcggaaatcctctcactcgtttgtggtgaaagcttgggtcaagcgttatacgagctggtgctg
    cgactgaca (97)
exon10:aatcactgcactcagctacgaaagcatgcttccactattccaagctcgggggagaatgcacatggaccaatcgggctcctgctg
    t (88)
exon11:gttgtcgccaaaacccttgactctgattggacactttttgctgcccctgtacgctatatacttcagcttcaagcttgaatcctg
    gctcacaataaccacagcactggtcaacagcgcagccatcctgtacagagcctgtgccgtcatggttctcctcataatactcggagtt
    acagtacctgggtttac (190)

```

(b) ゼブラフィッシュ *Danio rerio* *Squalene epoxidase* の各 exon におけるアミノ酸配列 (557 アミノ酸残基)

```

exon1:MWTFLLGIASF TYIYKKDAL LSSATTELLV AAVLCVSVGL IITCFSLRGQ TLKAPQFLQM PFQLLMTFSP TKELFSASST
    DSCSRSSKK (89)
exon2:RDV KRKRATDSQC SPTD DPEVIVGAG VLGSAMAVAL ARDGRVTVV GRDMKEPDRI VGELLQPGGF RVLKDLGLE (76)
exon3:(G)TVEGLDAH VVNGYVIHDM ESRTEVEIPY PQQENCIQI GRAFHHRGFI MGLRRAALAE (59)
exon4:(N)VKFVEGTV TSLEEEEGCV TGIQYREKDT GMLK (33)
exon5:EVRAPL TIVADGCFSK FRKNLVSGKV RVSSHFVGC I MK (38)
exon6:DPPQFKAH HAELVLADPS PVLIIQIISN ETRVLVDIRG EMPRDLMOYM TDKIYPQLP (57)
exon7:(E)HLKEPFMLAL QNDRLRTPA SFLPPAPV NK Q (32)
exon8:(G)VLLLGDAY NMRHPLTGGG MSVVLNDVSI WRDLKLNIPD LYDNTALLQ (48)
exon9:AKKKFHWERK SSHSFFVNVL TQALYELFAA TD (32)
exon10:(K)PLHQLRKA CFHYFKLGGE CINGPIGLLS (29)
exon11:(V)LSPKPLTLI GHFFAVALYA IYFSFKSESW LTIIPRALVNSA AILYRACAV MFPLKYSELQ YLVY (64)

```

図2 ゼブラフィッシュ *Danio rerio* *Squalene epoxidase* 遺伝子構造における exon 塩基配列とアミノ酸配列

() 内のアミノ酸一文字記号は、前の exon とで分断された遺伝暗号によるアミノ酸残基を示す。その場合、分断の後の exon でのアミノ酸残基としている。() 内の数字は、塩基対数またはアミノ酸残基数を表す。

表1 Blastp による BLAST 検索でヒットしたメダカゲノムデータベースのリード

ゼブラフィッシュ <i>Danio rerio</i>	ヒト <i>Homo sapiens</i>	ゼブラフィッシュ <i>Danio rerio</i>	ヒト <i>Homo sapiens</i>
exon1	exon1	exon8	exon8
olli7i01,olgi57k13 olli53i22,olli49o08 olli21f15,olli5p05 olli3e15,olli39d02 olli4i22,olli43f14 olli19p24,olbr62d24 olbr62j03,olea43b19	olli7i01,olgi57k13 olli53i22,olli49o08 olli21f15,olli5p05 olli3e15,olli39d02 olli4i22,olli43f14 olli19p24	rolec6d23,rolec3i21 rolova17i15,rolli52o19 rolki50n13,rolte32l22 rolbr9n16,rolbr32i09 rolli4i22	rolec6d23,rolec3i21 rolova17i15,rolli52o19 rolki50n13,rolte32l22 rolbr9n16,rolbr32i09
exon2	exon2	exon9	exon9
olgi57k13,olli53i22 olli49o08,olli43f14 olli4i22,olli21f15 olli7i01,olli39d02 olli5p05,olli3e15 olli19p24	olgi57k13,olli53i22 olli49o08,olli43f14 olli4i22,olli21f15 olli7i01,olli39d02 olli5p05,olli3e15 olli19p24	rolli4i22,rolli3e15 rolli53i22,rolli5p05 rolli33j03,rolli7i01 rolli49o08,rolli19p24 rolli43f14,rolgi57k13	rolli4i22,rolli3e15 rolli53i22,rolli5p05 rolli33j03,rolli7i01 rolli49o08,rolli19p24 rolli43f14,rolgi57k13
exon3	exon3	exon10	exon10
olli4i22,olgi57k13 olli49o08,olte23l08 olte25e09	olli4i22,olgi57k13 olli49o08	rolli3e15,rolli49o08 rolli33j03,rolli7i01 rolli19p24,rolli5p05 rolli53i22,rolli43f14 rolli4i22,rolli39d02 rolgi57k13	rolli3e15,rolli49o08 rolli33j03,rolli7i01 rolli19p24,rolli5p05 rolli53i22,rolli43f14 rolli4i22,rolli39d02 rolgi57k13
exon4	exon4	exon11	exon11
No hit data	No hit data	rolli39d02,rolli7i01 rolli21f15,rolli3e15 rolli49o08,rolli4i22 rolli5p05,rolgi57k13 rolli53i22,rolli19p24 rolli43f14,rolli33j03 rolec54e04,rolki8a16 rolki9f10	rolli39d02,rolli7i01 rolli21f15,rolli3e15 rolli49o08,rolli4i22 rolli5p05,rolgi57k13 rolli53i22,rolli19p24 rolli43f14,rolli33j03 rolec54e04,rolki8a16 rolki9f10
exon5	exon5		
No hit data	No hit data		
exon6	exon6		
No hit data	No hit data		
exon7	exon7		
No hit data	No hit data		

TTTCTAGTCATGAAATAATCATCAGCATCTCAGCTGTCAGGGAGGTTATAGAACCGCAGGTCCAGGGAGCGTTTTTTACATTTATGACAAGTTT
AACCAGAAAACCGCTCCGACCAACCTCATTATTTTCGATTATTTATTTACTTTCTGTTGAGAAGCAGCAGTTGATCGTCTATGTGCTTTTTTA
AGAAAATATTTAAAGATGTGGACTTTTCTAGGAGTGGCACTTTTACATATTTGTATAAAAAATCCAACCTGGTTCTGTCTTACATCCCAAAG
GAGCTGCTGATGATCACAGGCTTTGTTTGTGGCAGCTGGGCTGCTGCTGCTTATGTACAGTTCTTCCACGGCCAGAAGCTCACAGTCTTCCAGG
TGGTTCAATTTAAGCTTTACAGCGCTTGTCTTCTTACCCTGGTGAACCACTTAGTCCCTGAAAAAGCCCCACAGAGGAGCAGCAGCGGGAGAA
CAGCAGGAAAAAGGTAAGAGGCAGAGCTGTGGCAGTCAAAAGAAGAAAAACCAACAACCCCAAAGCTGACTGTCTGTTGTTTTAGATTTTTA
ATGATTTAAATCATTTGTGACATGTATTGATGTGAGCAGAAAGGATACATGCTTTTATGGATCTGTTTACAGAGTAAAGAGGAGCAGCAGAGAAG
AAACGGAGTCTGCCAGCTGAACCTGTTCTTCAAGTCTGTGGGCCCTGTGTGCCCGGAGCCGGAGCCGGATGTGGTGCATCGTTGGGCGCGGGT
GCTGGGCTCGGCCCTGTCCGTGGTGTGCTGGCCAGGACGGGAGGAGGGTACCGGTGGTGGAGAGAGACCTGAAGGAGCCGGACAGGATCGTGGG
GAGCTGCTGCAGCCCGGTGGATTAGAGCTCTGAGGGATCTGGGACTGGAAGGTGAGCAAAAAGCGGGGCTGGAGGTGATGAGGTCATGTGAG
GTGGCGTTTCACTCCACTTTTTGTTGAAAGTGTCAATGCAATCAGAAAACCATGAGGCAGAGGGAAAGCCCTTGATTAGCATCAACAAGATAAGG
GCTTTTTGTTATTAATAATGTTAATTTAATGTTTATTAATGTTTATTAATAGCTGTAGGATTAATAAAAAACCTGTTACTACTAACAGAACTTCCCTCCATGAA
106x6+38=636+38=674 deleted

TTGGACAGTGAAGATCAAGTTTAAAGATAAAAAATGCTTGTTTTTCTCTCAGATTTTCGTGTTGGAACATTTCCCACTGGACTGTAGAAGTGC
AACAAACACTTTTTGGAATAAAAAAACCAACCTTGTGAGTCTTTAACTTTGCGTGGAGTAACTGTCTGTGTTTCCCTCAGCTTTCAGTGGAG
AGGCTCGGATGCCAGCTGAACCTGTTGATCATCCATGACATGGAGAGCAGCCAGCTGGAGATCCCGTACCAGAGGAGGAGGGGGT
GTCAGTGTGGGCGCGCTTCCATCAGGTCGATTCAATTTAGGCTAAGGAGAGCCGTCATGGCCGAAAAAAAGTACGGAGGCACCTTTCTGT
ACGATATAACAGTGTGTTGAATACAGTAAAGTGAATACACTGATGATTAAAGTGGCGTCTCTGCACGGTCCGAGGTTGGTGTCCATGTTTC
TAACGGTCAACCTTGCCTTGTGTCAGGCTCATAGAAGGCACCGTGAACAGTTTACTGGAGGAGGAGCGGCTGTGTGACTGGAGTCCAGTAC
AGAGCAAAAGTGGCGGAGCTGAAGGTGAGCGCGGGGAACCTGCGCTTAAGCAGAGGAGAGGCTGCTTTTAGCAGGGTGCCTGTGACTGTAATGC
TGCCGTGAAACTGTCCCGCAGGAAAGATCCACGCGCGCTCACCGTGGTGGCCGACGGCTGCTTCTCCAAGTTCAGGAAGGCCGCTTCCCG
GGGCAAGACAACCTCACACTTTATCGGATGCTGATGAAGGTAGGCAGAGGAAAGCATCCAAATGCAGACTGAAACTGAAATGACTGGTTTAA
AAGATAAGATTCTATTGATGAGCAAAAGTGAACCTAAACTGTTAAATGCAGCCAAATTTAGATCAATTAAGGATTCCTTTTTTCCAACCCAGT
212+14=226 deleted

TGATACTGTGATACAGATAATCCAATAATAATAATCTGAAAGAGACTTTCTCAGAAAAGGATTATTTACTGTGAACTGATGTTTATTTGAGA
TCAAACCTGGCAGAAATGAATGAGCTAAAAGAGGAGAGAGACTCCACAGTGGATGTGCTGTCCAACGTCCTAAATGTCCACTGTTGTGCTGCCCC
CCCACCCAGACTGTCCGAGTTCAAAGCCAACCATCGCGAGCTGGTGTGGCCGACCCAGCCCGGTGTGATCTACCAGATCTCCTCCTCTC
ACACCCGGGTGCTGGTGGAGATCAGGGGGAGCTGCCTGCCTGACATGGAGAGGAGGCTGCTTTTAGCAGGGTGCCTGTGACTGTAATGC
AAACCAGACCGTTGGTGTGATCTGCGGTTAGATCGGATTTGTGTAACCGTAGATGTCCAGCCGCTTTAGCTGGAGAAGAAAAGCTGCAGTC
106+60=166 deleted

GTCAGCCCGCTGATGCATTGAGAAAACACTCAGACGCCCAAGAAAAGCTGTTGGTCCCAGCATGCACCTGCAGGGTTTCCCTCCAGACCAGTGA
CAGGCGTCTCCTCATGTTCTTTCCCTCCAGAGCACCTGAAAGAGCCCTTCATGGTGGGCTGCAGAACGACAGACTCCGGTCCATGCCCGAG
CAGCTTCCCTCCCTCCCTCCCGCTCAACAAACAGGTAAGTGGTGGTGGAGCCGTCGAGATGTTGAGGCAATATCTGTGCTGGCATTTT
GAACAATATAGTATAACCTTAAAGTTTGAACAGGAAGTCTGTTTTTTTTGCACGCAAGACCAAGTACCATTTCATTCATTTTGCAGCGCTGCT
GCTGTTGGGTGATGCCTACAACATGAGGCACCTCTGACGGGGGAGGGATGAGTGTCCGCTCAATGACGTCCGGATCTGGAGGGGTCTGCTG
AAGAACATCCAGATGCTGACGATGACTCCGCTGCTGAGGTTGCTCCCAACATGACTGTTCTCACAGGAAACCATAAACTGCATCCATC
TCTCTGGAATCATCAAAAATTTGGGCATAAAATGCATAATCCTTCAATACCTTCTGGACATACCTGCGGTTTTACGCCGTTTTTCCCCCTCT
TTTTCTCAATAGGCAAGAAAAAATTCACCTGGGAACGCAAGTCAATCCATCTTTTTGTGGTGAACGTGTTGGCGCAGGCTCTGTAGAGCTGTT
TTCAGCCACCGACAATGAGTCCCGTCTGCCTCTGCGGCCACAGACAGACAGAAAAAGAGTGTGGTCTCATCTCTGTTTGTCTGACA
GGTTCTCTTCATGACTCAGAAAGCCTGCTCCAGTACTTCAAGCTGGGTGGAGAGTATTTCTGGGCCATCGGACTTCTCTGTTGTAAG
CATCAATCAGTATCTTATCGCTTCACTCCCTCGGGAGCTGCTCAGCTTTGACTTTCCAGGAGTCCAAAAGGCAAAAAGCCTCATTTTTGCAT
AACAGAAGAACAAGTTAGGGAATT
(106x13+26=1060+318+26=1378+26=1404 deleted)

TCTTTTCCCTCAAAGTGTGTTAAAAAGGTGATCTAGACATAAAAAACGATCATGTTATCTGCATCATCTGAGGGTAGAAAAGCCTTTTCCA
GATTCTAAAATGAGCTTTTAGCTCTGAGTGAATGTTTGTGTTGAAGAAGTTTTTACAGCCTTCTTTTTTCTATCCACAATATCTTTTTTTG
TTTTGTTTCTAGCTGACACCGAAGCCCATGACCTCATTTGGACACTTCTTTGCGGTGGCGCTCTACGCGATCTACTTCGCTTTAAGTCTGAA
TCCTGGAGACCAACCTCGAGCTGTGTTCAAGAGCGCGCCATCCTGTACCAGCGTGCACGGTCATGTTCCCGCTCATCTACTCTGAACCTCA
AACACCTGCTTTAGTACTGAGGAGCCGAGCCACGGACGACAAAGGTTTGTGGTCTGAGACCCCTGGCAAAAATGTATCGCTTGGAGGGG
(378+60=438 deleted)

図3 メダカ *Oryzias latipes* ゲノム DNA における *Squalene epoxidase* exon 構造の推定

第16番染色体 (chromosome.16) の 2609251 ~ 2616200 の塩基配列に相当する。黒に白抜き文字が推定 exon で、左上の5'から第1~第11の順。アンダーライン (TGA) は第11exonの直後の終止コドン。

表2 メダカ *Oryzias latipes Squalene epoxidase* 各 exon の塩基配列と推定アミノ酸配列のまとめ

	塩基配列	推定アミノ酸配列
exon1	ATGTGGACTTTTCTAGGAGTGGCAACTTTTACATATTTGTATAAAAAATCCAAC CTGGTTCTGTCTTACATCCCAAAGGAGCTGCTGATGATCACAGGCTTGTTTGTG GCAGCTGGGCTGCTGCTGTCTTATGTCAGGTTCTTCCACGGCCAGAAGCTCACA GTCTTCCAGGTGGTTTCACTTAACCTTTCAGCGCCTTGTCTTCTTACCCTGGTG AACCACCTAGTCCCTGAAAAAGCCCCACAGAGGAGCAGCACGGCGGAGAACAGC AGGAAAAAG (279bp)	MWTFVLGVATFTYLYK KSNLVLVSYIPKELLM ITGLFVAAGLLLSYV RFFHGQKLTVFQVVH LTFALSFLPLVNH VPEKAPQRSSTAENS RKK (93)
intron	153bp	
exon2	AGTAAGAGGAGACGAGCAGAAGAAACGGAGTCTGCAGCTGAACCCTGTTCTTCA GTCTGTGGGGCTCTGTGGCCCCGAGCCGGAGCCGGATGTGGTGATCGTTGGG GCCGGGTGCTG GGCTCGGCCCTGTCCGTGGTGTGCTGGCCAGGGACGGGAGGAGGGTGACGGTGGTG GAGAGAGACCTGAAGGAGCCGGACAGGATCGTGGGGGAGCTGCTGCAGCCCGGT GGATTCAGAGCTCTGAGGGATCTGGGACTGGAAG (262 bp)	SKRRRAEETESAAEP CSSVCGASVAPEPEP DVVIVGAGVLGSALS VVLARDGRRVTVVER DLKEPDRIVGELLQP GGFRALRDLGLE (87)
intron	1082 bp	
exon3	CTTCAGTGGAAAGTCTGGATGCCACCAGGTGAACGGCTATGTCATCCATGACA TGGAGAGCAGCACCCAGGTGGAGATCCCGTACCAGAGGAGGAGGGGCGTGTCC AGTGTGGGCGCGCTTCCATCACGGTGCATTATTCAGGCCTAAGGAGAGCCG TCATGGCCGAAAAAAA (178bp)	(A) SVEGLDAHQVNG YVIHDMESSTQVEIP YPEEEGRVQCGRAPH HGRFISGLRRAVMAE K (59)
intron	138 bp	
exon4	TGTCAGGCTCATAGAAGGCACCGTGAACAGTTTACTGGAGGAGGACGGCTGTGT GACTGGAGTCCAGTACAGAGACAAAGACAGCGGAGAGCTGAAG (97bp)	(N) VRLIEGTVNSLL EEDGCVTVQYRDKD SGELK (33)
intron	88 bp	
exon5	GGAAAGATCCACGCCGCGCTCACCGTGGTGGCCGACGGCTGCTTCTCCAAGTTC AGGAAGAGCCTCGTTCCCGGGGCCAAGACAACCTCACACTTATCGGATGTCTG ATGAAG (114bp)	GKIHAALTVVADGCF SKFRKSLVPGAKTTS HFIGCLMK (38)
intron	570 bp	
exon6	GACTGTCCGAGTTCAAAGCCAACCATGCGGAGCTGGTGTGGCCGACCCAGC CCGGTGTGATCTACCAGATCTCCTCCTCTCACACCCGGGTGCTGGTGGACATC AGGGGGGAGCTGCCTCGCAACCTGTCCGAGTACATGGCGGAAAAAATACACCCA CAGCTGCCAG (172 bp)	DCPQFKANHAELVLA DPSPLIYQISSHT RVLVDIRGELPRNLS EYMAEKIHPQLP (57)
intron	394 bp	
exon7	AGCACCTGAAAAGAGCCCTTCATGGTGGCGCTGCAGAACGACAGACTCCGGTCCA TGCCAGCAGCTTCTCCTCCCTCCCTCCCCGTCACAAACCAG (96 bp)	(E) HLKEPFMVALQN DRLRSMPSFLPSP VNKP (32)
intron	146 bp	
exon8	GCGTGTGCTGTGTGGGTGATGCCTACAACATGAGGCACCTCTGACGGGGGAG GGATGAGTGTGCGCCCTCAATGACGTCCGGATCTGGAGGGGTCTGCTGAAGAACA TCCAGATCTGTACGATGACTCCGCCCTGCTGCAG (143 bp)	(G) VLLLGDAYNMRH PLTGGGMSVALNDVR IWRGLLNIPDLYDD SALLQ (48)
intron	157 bp	
exon9	GCAAAGAAAAAATCCACTGGGAACGCAAGTCATCCCATTTTGTGGTGAAC GTGTTGGCGCAGGCTCTGTACGAGCTGTTTTAGCCACCGACA (97 bp)	AKKKFHWERKSSHSF VVNLAQALYELFSA TD (32)
intron	81 bp	
exon10	GTTCTCTTCATGAGCTCAGAAAGCCTGCTTCCAGTACTTCAAGCTGGGTGGAG AGTGTATTTCTGGGCCATCGGACTTCTCTCTGT (88 bp)	(S) SLHELKACFQY FKLGGECISGPIGLL S (29)
intron	1703 bp	
exon11	GCTGACACCGAAGCCCATGACCCCTATTGGACACTTCTTTGCGGTGGCGCTCTA CGCGATCTACTTCGCCTTTAAGTCTGAATCCTGGAGACCAAACCTCGAGCTGT GTTCAAGAGCGCGCCATCCTGTACCGAGCGTGCACGGTCAATGTTCCCGCTCAT CTACTCTGAACTCAAACACCTGCTTTAC (190 bp)	(V) LTPKPMTLIGHF FAVALYAIYFAFKSE SWSTKPRAVFKSAAI LYRACTVMFPLIYSE LKHLLY (64)

※前 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 () でとじて、後の exon の最初に表示している。

```

exon1
O 1: MWTFLGVATFTYLYKSNLVLSYIPKELLMITGLFVAAGLLLSYVRFHFGKLTVFQVHLLTFSALSFLPLVNHVPEKAPQRSSTAENS
RKK
D 1: MWTFLGIASFTYIYKCDALSSATTELLVAAVLCVSVGLLIT-CFSLRGQTLKAPQFLQMPFQLMTFSPKTEFSA-----SSTDSCS
RSSKQRD
H 1: MWTFLGIATFTYFYKFGDFITLANREVLVLCVLLVLSLGLVLSYRCHRNRGGLGRQOSGSOEALFSDILSGLPFI GFFWAKSPPESENK
EQLE
exon2
O 94: SKRRRAEETESAAEPCSSVCGASVAPEPEPDVVI VAGVLSALS SVVLARDGRRVTVVERDLKEPDRIVGELLQPGGFALRDLGLE
D 92: VKRKRRA--TDS-----QCSPTEDDEVIIVGAGVLSAMAAVLARDGRRVTVVG RDMKEPDRIVGELLQPGGFVTKDLGLE
H 95: ARRRRKGTNISETSLIGTAACTSTSSQNDDEVIIVGAGVLSALAAVLSRDGRKVTVI ERDLKEPDRIVGEFLQPGGYHV KDLGLG
exon3
O 181: ASVEGLDAHQVNGYVIHDMESSTQVEIYPPE-EEGRVQCGRAFHHGRFISGLRRAVMAEK
D 166: GTVEGLDAHVVNGYVIHDMESRTEVEIYPPEQ-ENCIQGGRAFHHGRFIMGLRRAALAE P
H 182: DTVEGLDAQVNGYMIHDQESKSEVQIPYPLSENNOVQSGRAFHHGRFIMSLRRAAMAE P
exon4
O 240: NVRLIEGTVNSLLEEDGCVTVGVYRDKDSGELK
D 225: NVKFEVETVTSLEEEEGCVTGTQYREKDTGMLK
H 242: NAKFIEGVVLOLLEEDDVVMGVQYKDKETGDIK
exon5
O 273: GKIHAALTIVVADGCFSKFRKSLVPGAK--TTSHFICLMLK
D 258: -EVRAPLTIIVADGCFSKFRKNLVSGLKVRVSSHVGCIMK
H 275: -ELHAPLTVVADGLFSKFRKSLVSN-KVSVSSHVGVGLMK
exon6
O 311: DCPQFKANHAELVLADSPVLIYQISSSHTRVLVDIRGELPRNLSEYMAEKIHPQLP
D 296: DSQPFKAHHAELVLADSPVLIYQISSNETRVLVDIRGEMPRDLMQYMTDKIYPOLP
H 313: NADPQFKANHAELLANPSPVLIYQISSSETRVLVDIRGEMPRNLR EYMV EKIYPCIP
exon7
O 368: EHLKEPFMVALQNDRLRSMPSSFLPPSPV NKP
D 353: EHLKEPFMALQNDRLRTPASFLPPAPV NKP
H 370: DHLKEPFLEATDNSHLRSMPASFLPPSSVKKR
exon8
O 400: GVLLLGDAYNMRHPLTGGGMSVALNDVRIWRGLLKNIPDLYDD SALLQ
D 385: GVLLLGDAYNMRHPLTGGGMSVVLNDVSLWRDL LKNIPDLYDNTALLO
H 402: GVLLLGDAYNMRHPLTGGGMTVAFKDIKLWRKLLKGI PDLYDDAAIFE
exon9
O 448: AKKKFHWERKSSHSFVVNVLAQALYELFSATD
D 433: AKKKFHWERKSSHSFVVNVLTQALYELFAATD
H 450: AKKSFYWARKTSHSFVVNII LAQALYELFSATD
exon10
O 480: SSLHELKACFQYFKLGGECISGPIGLLS
D 465: KSLHQLRKACFHYFKLGGECINGPIGLLS
H 482: DSLHQLRKACFHYFKLGGECVAGFVGLLS
exon11
O 509: VLTPKPMTLIGHFFAVALYAIYFAFKSESWSTKPRAVFKSAAILYRACTVMFPLIYSELKHLLY
D 494: VLSFKPLTLIGHFFAVALYAIYSEFKSESWLTI PRALVNSAAILYRACAVMFPLIYSELYLVY
H 511: VLSFNPLVLIGHFFAVALYAVYCFKSEFWITKPRALLSSGAVLYKACSVIFPLIYSEMKYVMVH

```

図4 メダカ *Oryzias latipes* Squalene epoxidase 推定アミノ酸配列の比較

数字はアミノ酸の N 末端からの残基数。O: *Oryzias latipes* (メダカ), D: *Danio rerio* (ゼブラフィッシュ), H: *Homo sapiens* (ヒト)。黒に白抜き文字はメダカと同じアミノ酸, ■はメダカと類似した性質のアミノ酸, -はギャップ。ここでは, メダカの exon1 ~ exon11 に合わせて配列させている。