

論文の内容の要旨

論文題目 Construction of comprehensive genomic resources for medaka, *Oryzias latipes*
-Large-scale EST analysis and genome-wide phylogenetic study of the medaka
and its relatives-

(メダカゲノムリソースの構築

—大規模 EST 解析とメダカおよび近縁種のゲノムワイドな系統解析—)

氏 名 成田 貴則

現在、メダカやゼブラフィッシュのような小型魚類は生物学のさまざまな分野でモデル動物として広く使われている。両種のドラフトゲノム配列は UT ゲノムブラウザや Ensembl ゲノムブラウザ上にも公開され、すでに多くの研究者に利用されている。人為的な突然変異体誘発実験も行われ、ともに 1,000 種類を超える突然変異体が収集され、原因遺伝子の同定が精力的におこなわれている。

(1) EST 解析によるメダカゲノムリソースの作成

しかし私がこの研究を開始した 2000 年当時、メダカで利用可能なゲノム資源は Hd-rR および Cab 系統の BAC ライブラリー、ゲノム全体をカバーする連鎖地図及び千件ほどの EST 配列だけであった。我々の研究室でもメダカを脊椎動物のモデルとして、初期発生に異常を示す突然変異体のスクリーニングを開始していた。しかし、ゲノムリソースが不足していたため突然変異体の表現型解析やポジショナルクローニングを行うことはかなり困難をとまなう作業であった。このような現状を打開するため、私はまずメダカゲノムリソースの整備を行うことが重要であると考えた。EST 解析は大量かつ高速にゲノム情報を収集することができる方法として知られている。そこで私は初期発生と器官形成にターゲットを定めて EST 解析をおこなった。まず異なる 3 つの発生ステージ (ステージ 23 (体節形成期)、ステージ 35 (器官形成期) およびステージ 40

(稚魚期)の胚より 5 種類の cDNA ライブラリーを作製した。それぞれのライブラリーから 1 万クローン以上をランダムに選び、両端の塩基配列を決定した。その結果合計で 132,082 クローンの塩基配列を決定した。3'末端配列の相同性を用いて分類を行ったところ、12,429 の異なった塩基配列をもつクラスターに分かれた。脊椎動物の遺伝子数は 20,000-30,000 と考えられるため、いまだ飽和には達していないと考えられるが、少なくとも初期胚で発現する遺伝子の多くは網羅されており、現在でも多くの発生異常変異体の解析に有効なリソースを提供している。次に今回得られた大量の EST 配列を利用して、同時に多くの遺伝子の発現を解析できるマイクロアレイを作製した。8,092 クラスターの 3'末端配列中から 60 塩基の領域を選択し、DNA を人工合成後スライドガラス上にスポットした。EST 解析に用いた胚と同じステージに由来する RNA を用いて発現量を解析したところ、同一クラスター内に含まれるクローン数とマイクロアレイによる発現量は高い相関を示した。このことから、このマイクロアレイは有効に機能することが示された。また、発生とともにその発現が変動することが分かっている I-FABP 遺伝子の発現量を調べたところ、その変化を定量することができた。このマイクロアレイは、現在では遺伝子発現変化を調べる様々な研究に応用されている。

(2) メダカ及びその近縁種のゲノム情報の収集

私がおこなった EST 解析の後、2002 年からメダカゲノムプロジェクトが開始され、私も主要なメンバーとして参加した。この研究では南日本集団由来の近交系である Hd-rR 系統のゲノム DNA を用いて配列決定が行われた。Hd-rR 系統が選ばれた理由は、近交系であるため系統内の多型性がないこと、主な突然変異体スクリーニングプロジェクトで南日本集団由来の系統を用いていたことがあげられる。このプロジェクトでは、地域集団間の多型を調べ、その情報から高密度連鎖地図をつくることを目的として、北日本集団由来の近交系である HNI 系統の塩基配列決定も行った。メダカの南日本集団と北日本集団間は 2-3%の多型があることがすでに知られていたことから、両者を比較することで全ゲノム領域において多数の多型が見つかることが予想された。実際に解析の結果、1,600 万の SNPs と 280 万の挿入/欠失の多型を見つけることができた。この SNPs 情報を用いて全ゲノム領域をカバーする 2,500 以上の DNA マーカーをもつ高密度遺伝地図を作成し、ゲノムアセンブラーが生成した塩基配列 scaffold をこの連鎖地図上にならべることでドラフトレベルのゲノム解析を完了することができた (論文改訂中)。

私は、このようにして大量に得られた Hd-rR (南日本集団) と HNI (北日本集団) のゲノム情報を用いることで、メダカ系統内の多様性をさらに知ることができるのではないかと考えた。これまでのメダカ種内の系統関係はアロザイム、ミトコンドリア DNA 等を用いて解析され

てきた。その結果、メダカ自然集団は4つの地域集団（北日本集団、南日本集団、東韓集団、中国—西韓集団）に分けられることが知られていた。ゲノム解析によって南日本集団由来の Hd-rR と北日本集団由来の HNI 間のゲノム配列には平均 3.4%の相違があることは分かったが、その他の集団間の比較は行われていなかった。そこで、私はメダカのもつ24対の染色体から4ヶ所ずつ、合計 96 の領域を増幅することができるプライマーを用いて、各地域集団由来のゲノム DNA を鋳型として PCR を行い、ダイレクトシーケンス法でそれらの塩基配列を決定した。最終的に、4つの地域集団に属する 12 系統から 47 組の配列（1 系統あたり約 20k bp の塩基配列をゲノムよりサンプリング）を決定することができた。これらの配列を多重整列したところ 948 の SNP サイトを同定した。このデータを用いて近隣結合法により系統樹を作成したところ大きく4つの集団にわかれた。各集団の単系統性は高いブートストラップ値（100%）によって支持された。今回の解析で認められた4つの単系統集団はミトコンドリア DNA の塩基配列に基づいて分類された南日本集団、北日本集団、中国—西韓集団及び東韓集団と完全に一致した。多型サイトの分布を調べたところ集団内多型が小さいのは北日本集団と東韓集団であった。中国—西韓集団と南日本集団は比較的大きな遺伝的多様性をもっていた。核遺伝子とミトコンドリア遺伝子それぞれから得られた系統樹は概ね同一の樹型を示したが、中国—西韓集団内の樹型だけは異なっていた。ミトコンドリアでは Kunming 系統が最初に分岐しているのに対し、核遺伝子では Taiwan 系統が最初に分岐していた。この結果は Kunming 系統でのミトコンドリア移入の可能性を示唆している。これは系統解析にミトコンドリアとともに核遺伝子を用いることの重要性を示した一例でもある。

メダカのモデル動物としての優位性は、飼育しやすい、毎日採卵可能、胚が透明、短い世代時間、トランスジェニック技術が確立している、ゲノム配列が決定されているなどであろう。しかし、このような特徴は他のモデル動物にもあてはまる。ではメダカがマウスやゼブラフィッシュなどの他のモデル動物と大きく異なる点はなんだろうか。私はメダカの自然集団がもつ大きな遺伝的多様性（メダカはいままで知られている脊椎動物では最大の種内変異をもつ）と近縁種が存在し、系統保存されていることであると考え。メダカとその近縁種はダツ目メダカ科（Adrianichthyidae）に含まれるが、いままでに 4 属 26 種が記載されている。これらの系統関係は、核型、アロザイムおよびミトコンドリア塩基配列によって解析されてきた。最新の研究では、メダカとその近縁種は3つのグループ（*latipes*, *javanicus* および *celebensis* グループ）に分けられることが示されている。これら 3 グループ内の単系統性は高いブートストラップ値（97-100%）によって支持されている。しかし、これら3つのグループ間の系統関係はいまだ十分には分かっていない。私は、今回設計したプライマーをメダカ近縁種に応用することによって、

メダカ近縁種のゲノム情報を収集するとともに、その塩基配列から 3 者の系統関係を明らかにすることを試みた。それぞれの種から抽出したゲノム DNA を鋳型として種内変異の解析に用いたプライマーによってそれぞれの種から相同な領域を増幅した。その後ダイレクトシーケンス法によって増幅した DNA 断片の塩基配列を決定した。得られた配列からエクソン領域を特定し、配列を結合して系統解析を行った。しかし、これら 3 グループ間の系統関係はいままでの報告と同程度の精度であった。その理由は、系統解析に十分量の相同なエクソン領域の塩基配列を、外群にもちいたミドリフグのゲノム情報から同定することが困難であったためである。しかし、核遺伝子配列を用いて得られた系統関係は、これまでの解析と矛盾する部分がないことから特定の系統間での大規模なミトコンドリア遺伝子移入などの現象は起こっていないと考えられる。今回の解析では、メダカを外群として、スラベシ島内に生息するメダカ近縁種 (*celebensis* グループに属する) の系統関係も解析した。*celebensis* グループ内の系統関係はグループ内の遺伝的分化が乏しいために十分な解像度が得られていなかった。私は約 13 k bp の塩基配列を用いて、*celebensis* グループ内の系統解析をおこなった。*X. oophorus* と *X. sarasinorum* は形態的に類似しているため、同じ属に分類されている。しかし、私の解析から *X. oophorus* と *X. sarasinorum* は単系統ではないことがあきらかとなった。この結果は両者で共通にみられる「孵化まで卵を腹部で保育する」という生殖様式は沖合を生息域として、回遊するという生態に適応したことによる収斂の結果であることが示唆された。今回の解析から 3 つのグループを区別することができる分子的な共有派生形質 (特定のグループがのみがもつ欠失/挿入) を多数同定することもできた。このことは塩基配列の類似性だけではなく、塩基の挿入/欠失という現象から見た場合にも 3 つのグループが単系統性を示すことを示唆している。興味深いことに 6 ヶ所の独立な欠失が *javanicus* と *celebensis* グループで共有されていた。この結果といままでに報告されている系統関係とを考えると *javanicus* グループと *celebensis* グループが共通祖先をもつことを示唆しているのかもしれない。

一連の研究によって私は EST を中心としてメダカゲノムリソースを整備することができた。またメダカゲノムから特定領域を増幅できるプライマーを用いて、南北日本集団のみではなく、東韓集団、中国—西韓集団を含むメダカ全体の遺伝的多様性の概要を示すことができた。またメダカ近縁種を特徴づける分子的な共有派生形質を多く見つけるとともに、いままで明らかでなかった *celebensis* グループ内の新たな系統関係を示すことができた。これらの情報はメダカとその近縁種をとともにもちいたモデル動物の確立に貢献すると考えている。