

高精度比較ゲノム地図の作成と、それに基づいた比較ゲノム構造解析研究

●藤山 秋佐夫¹⁾ ◆二階堂 雅人²⁾ ◆豊田 敦³⁾

1) 情報・システム研究機構国立情報学研究所 2) 東京工業大学大学院生命理工学研究科 3) 理化学研究所ゲノム科学総合研究センター

<研究の目的と進め方>

本研究では、我々がこれまでに進めてきた比較ゲノム研究を継続的に発展させ、霊長類から単孔類までの哺乳類ゲノムと硬骨魚類ゲノムを対象とした系統的かつ網羅的な比較解析を通して、ヒトへと向かう進化経路で起きた進化イベントをゲノムレベルでの変遷過程として解明することを目指す。研究の実施に当たっては、特に常染色体と性染色体の比較、テロメア領域、反復配列、シンテニー等のゲノムを構成する単位に着目したい。一方、種の形成に関する問題は、17年度の間中評価でも指摘されたように生物学全般にわたる基本的で重要な課題である。そこで本研究では、生物種の形成と多様化のモデルケースとしてアフリカ東部ビクトリア湖に生息する硬骨魚類であるシクリッドをとりあげ、近縁種間の比較ゲノム解析から、この種で起きた種分岐の機構解明を目指す。以上により、最終的には、脊索動物の出現から哺乳類、霊長類、ヒトへの進化という文脈の中で起きたと考えられるゲノムの構造変化過程の主要経路を明らかにしたい。そのため、これらの研究の基礎となるリソースの整備を支援班の協力を得ながら進めると共に、高精度地図の構築を進め、今後も重要性が増加することが予想される比較ゲノム解析のための研究基盤の整備も図る。研究組織としては、渡邊正勝班員の転出に伴い、平成19年度から新たに二階堂雅人を分担班員とした。

<2007年度の研究の当初計画>

2005年に、概要配列ではあるがチンパンジー全ゲノム配列が報告され、2006年には改訂版が公開された。また2007年には、アカゲザルゲノムの概要配列解析結果が論文発表になり、さらに霊長類以外の哺乳類ゲノムでは、ウシ、ブタ、オポッサム、ワラビーなどの概要配列決定が進むなど、比較ゲノム研究を取り巻く世界的な情勢が米国を中心に急速な展開を遂げている。このため、本研究も世界の研究情勢を見極め、あらためて戦略を立て直す時期に到達している。そのための方策の一つとして、研究項目B02に属し、霊長類ゲノムを直接の研究対象としている6グループの連携により、霊長類、ほ乳類ゲノムでの高度保存領域の解析、霊長類におけるサブテロメア領域の比較ゲノム解析、性染色体/常染色体比較解析を年度内に開始する計画とした。さらに、これらの計画を補完するため、豊田班員は、霊長類の大規模ゲノム構造解析、特にヒト21番染色体に相同(シンテニー)な領域の高精度な配列決定を行い、進化における染色体の構造変化を明らかにする目的で、オラウータンの染色体特異的 fosmid ライブラリを構築するとともに次世代シーケンサーを用いて領域特異的な高精度ゲノム比較を開始する。また、ヒトモデル動物として期待されているコモンマーマセットについては、ゲノム研究の基盤情報の構築を進める。テロメア領域の解析については、フィンガープリント作成による部位の特定や定量PCRによるコピー数の測定などの新しい手法を導入して解析を続行する。前述したように国際状況には厳しいものがあるが、相互に類似性の高い霊長類間でのゲノム比較では、高精度の解析が必要不可欠であり、我々が独自に開発してきたゲノム資源を有効に活用することで進化の過程におけるゲノム構造変化と機能に与えた影響の全貌を明らかにするこ

とで、国際的な寄与を果たしたい。

一方、東アフリカのビクトリア湖産シクリッドは、わずか12,000年という、進化的にみるとごく最近に種の放散が起きたと考えられており、DNAレベルでの多様性が著しく低いにも関わらず、形態的・生態的に多様な独立した種が多数存在している。二階堂班員は、種や形態を決定付ける因子は「DNAレベルでの違いが少ない」方が実験的に特定しやすいとの仮定のもとに研究を進め、遺伝子重複や、また、魚類の多くではいまだ明らかにされていない性決定の問題も含めて染色体レベルの比較解析を進める。そのため、まずは嗅覚受容体遺伝子の中から特にOR型遺伝子群とV2R型遺伝子群に着目して網羅的な単離を試みる。

<2007年度の成果>

霊長類を対象とした研究として、B02 斎藤班、植田班、那波班、黒木班との共同研究を新たに開始し、ヒト19番染色体長腕テロメア領域2Mbを対照領域とし、まずチンパンジーゲノムについて相同領域のクローン地図作成を開始した。ヒト19番染色体はテロメア末端までクローンコンティグが作成されていない、つまり配列情報が存在しないため、この領域を対象としたコンティグクローンの単離を進めている。さらに、斎藤班、植田班、黒木班との共同研究で、ゴリラゲノムテロメア領域の比較研究も開始し、全ゲノムフォスミドライブラリからのテロメア末端クローンの単離と細胞遺伝学的解析を進めつつある。霊長類、哺乳動物ゲノム解析材料として、支援班石田班員から供与された細胞株を用いてオランウータンBACライブラリの構築を進めた。また、オーストラリア国立大学Gravesグループとの共同研究も進めており、カンガルー全ゲノムのBACライブラリ構築、X染色体特異的フォスミドライブラリを構築した。また、基盤ゲノム菅野班の支援と東京医科歯科大学石野グループとの共同研究により、カンガルー精巣、卵巣、子宮膜、下垂体、脳から全長cDNAライブラリを構築した。豊田班員はオラウータンゲノムからセルソーターで単離したヒト21番染色体に相当する染色体DNAを増幅後、次世代シーケンサー(GS20)により配列決定を行った。その結果、2,152,082リードのうち約80%がアセンブリに使用でき、コンティグ数は約13,200(総塩基数:約14Mb)、そのうちの約95%がヒト21番染色体にマップ(平均identityは95.8%)された。当初計画していたフォスミドライブラリが作製できなかったため、増幅後のゲノムDNAを使用してサンガー法によるショットガンデータを加え、現在評価を行っているところである。コモンマーマセットについては、これまでに国内の大学・研究機関と共同で行った完全長cDNAライブラリの5'末端配列(5'EST)の結果を利用し、マーマセットの基盤情報の構築を目指して特にES細胞で発現している遺伝子の全長配列決定を進めている。

二階堂班員が担当しているシクリッドの多様化研究については、前年度に単離したOR遺伝子群の解析に引き続き、合計11個のOR遺伝子を単離することに成功した。これらのOR遺伝子のタンパク質コード領域について、ビクトリア湖産シクリッド間におけるDNA配列を比較したところ、ある1つの遺伝子に関して、種間で異なるアリル型が確認された。今後は、このOR遺伝

子のアレル頻度がシクリッドの種間における有意差の検定を行う予定である。V2R 遺伝子群に関しては、遺伝子クラスター領域を全てカバーする領域を8個のBAC cloneに絞り込むことに成功した。ただ非常に高頻度な反復配列を含む領域（おそらく20kbp以下でその領域にはV2R 遺伝子は存在していないと予想している）のみは未単離となっている。これらのBAC cloneについては全配列決定を終了した。

<国内外での成果の位置づけ>

2004年に完成したことになるヒトゲノムであるが、サブテロメア・テロメア領域に関する限りは、信頼性のあるクローン地図が作成されておらず、したがって日本の研究グループが構造決定に関与してきた一部の染色体を除き、構造情報は極めて不十分である。特にテロメア領域の構造不安定性が細胞老化やガン化との関連で問題になっていることから、この領域の正しい構造情報を決定することは極めて重要であると考えられる。また、サブテロメア領域の構造については、細胞遺伝学的研究からヒト、チンパンジー、ゴリラ、オランウータンの間でかなりの違いのあることが知られていたが、実際にチンパンジーとヒト染色体のテロメア領域の構造とを高精度配列を基に比較することにより、具体的な配列情報の違いが明らかになりつつある（B02 公募班、黒木陽子との共同研究）。米国では、オランウータン、コモンマーモセットともにゲノム配列決定が進められているが、我々の研究成果はオランウータンについては個体差間の多様性を、コモンマーモセットについては発現遺伝子の構造およびヒトとの違いを明らかにできると考えている。

モデル魚種（ゼブラフィッシュ、メダカ、フグ、トゲウオ等）以外における魚類嗅覚遺伝子群の解析は、ほとんど進んでいない。それはタンデムに重複している嗅覚遺伝子群の網羅的な塩基配列決定や、その解析が非常に困難なためである。今回、V2R 遺伝子群を包含するゲノム領域の単離に成功したが、続いてその配列の解析が終了すれば、それはモデル魚種以外では世界で最初の例になる。またシクリッドは一般的なモデル魚種と違い、遺伝子を単離した後の種間比較といった観点から非常に重要な生物である。これまで、嗅覚受容体遺伝子群の種間、集団間における大規模な比較研究はヒト・マウスに限られており、それも2007年に研究論文が出版されたばかりの新しい分野である。シクリッド野生集団において嗅覚受容体遺伝子にどれだけの個体差、種間差が存在しているのかに関しては多くの研究者の興味を集まるところとされている。

<達成できなかったこと、予想外の困難、その理由>

当初の予想以上にサブテロメア領域の解析が難航しており、未だにヒトゲノムの全テロメア領域のクローン化には至っていない。主要な原因はサブテロメア領域に染色体間で共有される領域が多いためにクローンの特定に手間取っていることと、おそらくはクローン化の時点でのバイアスと宿主大腸菌内での不安定化などが原因と思われる。また、ヒトゲノム解読コンソーシアムから提供されている配列には利用不可能なクローンやYACなどの本来この領域の解析には不向きなベクターを用いた結果がかなり多く含まれており、全般的な見直しが必要である。ヒト19番染色体と21番染色体霊長類間での比較を進めることでこの問題の一部を解決したい。ソーティングしたオランウータン染色体が量的に不十分であったため、本年度はまず配列決定を先行させ、fosmidライブラリ作成は次年度に行うように計画を変更した。

シクリッドについては、今回単離したV2R 遺伝子群が含まれるゲノム領域に、非常に高頻度な反復配列が存在し、その部分をカバーするBACクローンを単離できなかった。おそらく、その高頻度反復配列内には、目的とするV2R 遺伝子は存在しないと考えられるが、BACクローンによるコンティグを完成させるた

めには、この領域に由来するクローンの単離が今後解決すべき問題である。そのための対策としては、BACライブラリを使用せずに長距離PCR法で該当領域をゲノムから直接増幅させることや、別の制限酵素を使用したファージライブラリの併用をすることで、解決できると考えられる。

<今後の課題>

今年度の後半に4班の共同研究として開始した霊長類テロメア、サブテロメア領域の比較解析研究については、テロメア領域を含むクローンをライブラリから網羅するスクリーニングを続行すると共に、ヒト19番染色体長腕末端2 Mb に対する相同領域について、特にこの領域に遺伝子クラスターが存在する zinc finger タンパク群を中心に網羅的解析を行う。また、ヒトゲノムについては、全染色体のテロメア・サブテロメア領域について構造決定の完成を目指す。ゲノム構造決定については、次世代シーケンサーを使用して配列決定を加速させていく予定である。また、クジラ類は嗅覚受容体遺伝子だけではなく、味覚（クジラの舌には味蕾がないかわりに小孔が存在し、そこにある受容体が水中に溶けた化学物質を認識していると考えられている）や触覚（触覚毛）などの特徴的な適応進化を遂げているため、これについてもゲノム構造を明らかにしたいと考えている。ビクトリア湖シクリッド間における、嗅覚受容体遺伝子群の進化機構と併せて解析を進めたい。

<成果公表リスト>

1) 査読付き論文

1. 0801270031

Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, Yamada T, Nagayasu Y, Doi K, Kasai Y, Jindo T, Kobayashi D, Shimada A, Toyoda A, Kuroki Y, Fujiyama A, Sasaki T, Shimizu A, Asakawa S, Shimizu N, Hashimoto S, Yang J, Lee Y, Matsushima K, Sugano S, Sakaizumi M, Narita T, Ohishi K, Haga S, Ohta F, Nomoto H, Nogata K, Morishita T, Endo T, Shin-I T, Takeda H, Morishita S, Kohara Y.: The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714-719 (2007).

2. 0801270044

Ahsan B, Kobayashi D, Yamada T, Kasahara M, Sasaki S, Saito TL, Nagayasu Y, Doi K, Nakatani Y, Qu W, Jindo T, Shimada A, Naruse K, Toyoda A, Kuroki Y, Fujiyama A, Sasaki T, Shimizu A, Asakawa S, Shimizu N, Hashimoto S, Yang J, Lee Y, Matsushima K, Sugano S, Sakaizumi M, Narita T, Ohishi K, Haga S, Ohta F, Nomoto H, Nogata K, Morishita T, Endo T, Shin-I T, Takeda H, Kohara Y, Morishita S.: UTGB/medaka: genomic resource database for medaka biology. *Nucleic Acids Research* 36, D747-752 (2008)

3. 0801251617

Stefan A. Rensing, Daniel Lang, Andreas D. Zimmer, Astrid Terry, Asaf Salamov, Harris Shapiro, Tomoaki Nishiyama, Pierre-François Perroud, Erika Lindquist, Yasuko Kamisugi, Takako Tanahashi, Keiko Sakakibara, Tomomichi Fujita, Kazuko Oishi, Tadasu Shin-I, Yoko Kuroki, Atsushi Toyoda, Yutaka Suzuki, Shin-ichi Hashimoto, Kazuo Yamaguchi, Sumio Sugano, Yuji Kohara, Asao Fujiyama et al.: The genome of the moss *Physcomitrella patens* reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64-69 (2008)

比較ゲノム解析による進化・多様性ゲノム基盤の解明支援

●藤山 秋佐夫¹⁾ ◆岸野 洋久²⁾ ◆久原 哲³⁾ ◆堀 寛⁴⁾ ◆遠藤 秀紀⁵⁾ ◆石田 貴文⁶⁾
◆黒木 陽子⁷⁾ ◆豊田 敦⁷⁾ ◆笠原 雅弘⁸⁾ 鈴木 穰⁸⁾

1) 国立情報学研究所 2) 東京大学大学院農学生命科学研究科 3) 九州大学大学院農学研究院 4) 名古屋大学大学院理学研究科
5) 京都大学霊長類研究所 6) 東京大学大学院理学研究科 7) 理化学研究所ゲノム科学総合研究センター
8) 東京大学大学院新領域創成科学研究科

<目的>

支援班は、研究支援を目的に本特定領域の開始年次から新たに特定領域内に設定されたものだが、本支援班ではゲノム4領域全体を対象とした支援活動を行っている。比較ゲノム領域に設置された本支援班では、1) 特別な技術や設備が必要、2) 個別の研究班で対応するよりは領域全体として対応の方がふさわしい、3) 各分野のエキスパートの協力を必要、とする項目について、領域の研究推進に必要な研究支援を行う。本支援班の機能を有効に利用することにより、比較ゲノム研究に特徴的な最新の技術や設備へのアクセスと、各分野の専門家との密接な交流が可能になり、各班における研究の質と効率の向上が期待される。支援の実施に当たっては、支援要請者と十分に協議し、内容と必要性に応じて本支援班、基盤ゲノムの各班と支援班、もしくは適当な班員と連携して実際の支援を行う。また、当班は支援班であるため、支援に必要な技術開発は行うが研究は実施しない。

<2007年度の活動方針>

今年度の本班の活動は、昨年度と同じ方針を引き継いだ。

- (1) 支援対象は、ゲノム4領域全ての研究班とする。分担班員からの支援要請は研究代表を経由する。また、計画研究班については、領域設定時に組み込まれている課題もある。
- (2) 支援の案内は、領域班会議、全体班会議において口頭で行うと共に、メイルリストを利用し全班員に通知した。
- (3) 支援依頼は、本班代表者宛のメイルで随時受け付ける。
- (4) 各支援依頼は、本班班員と領域内総括班員で組織した比較ゲノム研究支援委員会で審査のうえ、支援担当班と協議し具体的な支援を実施する。
- (5) 支援に際しては、研究班に共同研究など適宜協力を要請する場合もある。
- (6) 通知内容は以下の通り。

領域2「比較ゲノム」には、「比較ゲノム解析による進化・多様性のゲノム基盤の解明支援」班が設けられており、細胞、ライブラリ等の研究用リソースの整備・作成支援、シーケンシングや配列解析などのゲノム基盤情報作成支援、マイクロアレイ実験等の実験技術に関する助言・支援を予定しております。昨年度は、BACライブラリ、フォスミドライブラリの作成とスクリーニング、(全長) cDNAライブラリ作成、シーケンシングと配列解析、マイクロアレイ講習会、材料の持ち込みによるアレイ作成支援等を他の支援班、「基盤ゲノム」各班との協力の下に実施し、相当な成果を上げることができています。支援班は特定領域研究にユニークな活動であり、班員の皆様には是非有効に活用いただいて成果につなげていただきたいと思います。つきましては、全領域の研究代表者を対象に、下記の要領で支援の希望についてお聞きしたいと思います。分担班員の先生には代表者からご連絡をいた

き、代表者経由でお申し込み下さい。

具体的な支援につきましては、国内研究者の収束状況、国際的なプレゼンス、予算状況などを考慮し、支援委員会で検討した後、基盤ゲノム各班とも相談した上で内容と順位を決定いたします。また、状況によりましては経費の一部負担、マンパワーの提供をお願いする場合もある他、適当な協力先をご紹介する場合がありますことをご承知おき下さい。

<2007年度の成果>

前年度からの継続分に加え、今年度は新たに以下の支援を実施した。

研究用リソース作成支援：

ミヤコグサ、ヒメツリガネゴケ、クラミドモナスについて、複数条件下でcDNAライブラリを作成（進行中）

Acytostelium粘菌の、BACライブラリ、フォスミドライブラリの作成（進行中）

ハブスメダカBACライブラリの作成（細胞株の準備中）

シーケンシングと配列解析支援（基盤ゲノム小原班、基盤ゲノム情報支援班、基盤ゲノム菅野班とも連携）

新型シーケンサによる植物ESTシーケンシング、転写開始点情報の大規模取得（シーケンサの準備中）

Acytostelium属粘菌cDNAシーケンシング（進行中）

Acytostelium粘菌のBAC&フォスミドクローンの末端配列解析（進行中）

ヒトRNA中に含まれるエディティング部位および修飾部位を網羅的に探索（シーケンサの準備中）

クジラBACコンティグの配列決定シクリッド嗅覚遺伝子を含むBACの配列決定

寄生性原虫cDNAの配列決定（進行中）

パンコムギESTシーケンシング

<次年度計画>

今年度、新たにIllumina/Solexaシーケンサ、レーザーダイセクション装置、定量PCR装置を導入し、稼働テストを行っている。これらの装置については平成20年度から一般支援に供する予定である。