

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⁾²⁾ ◆豊田 敦²⁾ ◆黒木 陽子³⁾

1) 国立情報学研究所 2) 国立遺伝学研究所 3) 理化学研究所

<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する。近年の次世代シーケンサの登場により、個体（系統）間や近縁種間のゲノム構造の違いやさまざまな環境に存在する微生物集団の組成を迅速に低コストで明らかにすることが可能となりつつある。そこで本研究では、次世代シーケンサによる配列解析システムを構築し、遺伝的多様性を明らかにすることも新たな目標に設定した。特にヒトでは遺伝的多様性がさまざまな疾患や体質、薬物応答に関与していることが示唆されており、今後テーラーメイド医療や予防医学を実現するにあたって、その基盤技術の構築が必要不可欠である。

20年度の研究組織としては、分担班員の入れ替えと研究協力者（西原秀典：東京工業大学）を追加した。

<2008年度の研究の当初計画>

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。こうした状況に対処するため、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昨年度から開始した。特に、細胞遺伝学的手法によってヒトと 3 種の類人猿染色体間で構造変異が確認されている染色体サブテロメア領域（ヒトとオランウータンとの間、ゴリラとチンパンジーとの間で見かけ上の類似性が高い）、なかでも代表的な遺伝子調節タンパクである Zn フィンガータンパクがクラスターを形成して存在する 19 番染色体長腕末端 2Mb の領域に着目し、霊長類間での比較ゲノム解析を徹底的に実施することで、代表的な遺伝子調節タンパクである Zn フィンガークラスター（と下流の遺伝子群）メンバーの増減とヒトの進化過程との関連を明らかにする計画である。

しかし一方で、現在公開されているヒトゲノムアセンブリには、5 番長腕、10 番短腕、11 番長短腕、18 番短腕、21 番長腕（以上は日本グループの寄与）と、X/Y 短腕以外には配列解読がテロメアまで到達した染色体は存在しない。そこでまず、ヒトゲノムのテロメア領域を対象に信頼性の高いライブラリからクローンを単離し、再シーケンシングによって構造を再確認する構造解析研究を並行的に実施し、この情報をもとに他の霊長類ゲノムからのクローン単離や構造比較を行い、ヒトに至る染色体の再構成プロセスを明らかにする。以上の研究には、本研究班以外に研究項目 B02 植田班、斎藤班、那波班が密接に協力して構造解析に当たる。さらに、これらの計画を補完するため、豊田班員と黒木班員は、霊長類の大規模ゲノム構造解析、特にヒト 21 番染色体と Y 染色体に相同（シンテニー）な領域の構造解析を進める。そのため、オランウータンの染色体特異的 fosmid ライブラリのスクリーニング、構造決定、純化染色体の次世代シーケンサによるゲノム比較を開始する。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。

また、個人ゲノム解析の基盤整備に向けて、ヒト由来の細胞株を用いて次世代シーケンサによる配列解析および構造変異の検出手法を構築する計画である。

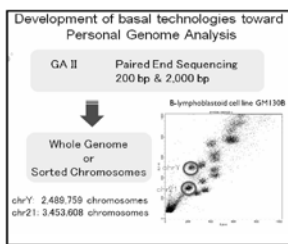
<2008年度の成果>

霊長類サブテロメア領域のゲノム解析：ヒト 19 番染色体長腕サブテロメア領域と、それに対応するチンパンジー、オランウータンのゲノム構造解析、ゴリラサブテロメア領域由来クローンの解析を行った。現行のヒト参照配列内にサブテロメア領域が含まれていなかったため、まず、この領域のクローン地図作製を行った。この領域に対応するチンパンジーとオランウータンのゲノム領域も解析を進めており、チンパンジーについては、ヒト 19 番長腕サブテロメア領域近傍のクローン地図を完成した。一方、ゴリラ全ゲノムフォスミドライブラリーから、サブテロメア領域由来のクローンを単離し、染色体上へマッピングしたところ、複数の染色体のサブテロメア領域に相同性を示すクローンが得られたため、これらのクローンの塩基配列決定を進めている。

哺乳類 Y 染色体の比較ゲノム解析：本年度は、昨年度から進めているヒト-オランウータン Y 染色体の比較クローン地図の高精度化と高被覆度化を目標に、finger printing によるクローンの整列化とゲノムライブラリーのスクリーニングを行った。今回、新たにオランウータン Y 染色体特異的フォスミドライブラリーを作成し、Y 染色体サブテロメア領域由来のクローンを単離するとともに、サブテロメア側からセントロメア側に向けてクローン地図の作製に着手した。これまでに、オランウータン Y 染色体の長腕と短腕、それぞれのサブテロメア候補クローンを得られており、これらのクローンをを用いた他の大型類人猿の染色体上の位置を確認するとともに、クローンの高精度塩基配列の決定を進めている。また、ヒト Y 染色体の遺伝子領域に対応するオランウータン Y 染色体のゲノム構造解析も、昨年に引き続き、解析を進めている。

次世代シーケンサを用いた染色体特化型ゲノム解析：次世代シーケンサを用いたパーソナルゲノム解析に向けた技術開発を目指し、ヒト 21 番染色体と Y 染色体の塩基配列決定を行った。リンパ芽球様細胞株 GM130B からセルソーターを用いて 21 番染色体（約 350 万本）を単離し、Genome Analyzer II を使用して決定した配列データ（約 9224 万リード）をヒト標準配列にマッピングするとともにアセンブリにより作成されたコンティグ配列との比較を行った結果、平均重複度 33.8x で約 36,600 カ所の SNP を同定できた。構造変化を視覚化する viewer を作製中（遺伝研・小原研究室と共同）である。なお、アセンブリについては、データ量が少ないためミスアセンブリが生じていることが示唆された。また、Y 染色体についてはソートできた染色体数が少ないためデータ解析が遅れており、配列決定を続行中である。

比較ゲノム構造進化研究として、新たに哺乳類が海洋環境へ適応する過程に起こったゲノム構造の変化に着目し、クジラ類



(オキゴンドウクジラ)のBACライブラリ作製とクローン地図の作成を行った。特に嗅覚関連遺伝子群領域とMHC領域(東海大・椎名班員との共同研究)に注目し、各領域に位置するBACクローンのスクリーニングと配列決定を行った。これまでに海生哺乳動物と陸生動物の大規模なゲノム比較の例はなく、進化の過程で起きたゲノム構造の変化を明らかにしたいと考えている。

<国内外での成果の位置づけ>

本研究では、染色体領域として常染色体である21番染色体と雄特有な染色体であるY染色体を解析対象とし、全ての染色体に共通に存在すると考えられているサブテロメア領域、その中でもゲノム構造解析が進んでいる21番染色体長腕、遺伝子密度が高くヒトとチンパンジーの脳における発現パターンに特異性があることが示唆されている19番染色体の長腕領域のサブテロメア領域に焦点を絞り、解析を行っている。Y染色体については、本研究と同様な哺乳類の比較ゲノム解析が米国のシーケンスセンターで進められているが、解析対象生物が異なっており、ゲノム構造や塩基配列のデータなど、解析結果はそれぞれの研究に有用であるが、直接的に競合することはない。また、タマウラビーのY染色体比較ゲノム解析は、オーストラリアの研究チームと共同で進めており、有袋類や単孔類のゲノム研究の専門家、性分化と生殖発生学の専門家と連携して解析を進めている。ゲノムの比較解析から明らかになったY染色体の機能分子は、有袋類の個体発生過程における機能解析が可能であり、ゲノム構造の変化と表現型の関連性を明らかにすることができると考えている。

1000人ゲノムプロジェクトと呼ばれる、個人のゲノム解析プロジェクトが進んでいる一方で、Y染色体のように反復配列が豊富な上に全ゲノムでハプロイドのゲノム領域は、配列データの解析が難しいことが危惧されている。そこで、本研究では、染色体特化型のリシーケンス法による解析系の立ち上げを進めている。ソーティングした染色体を用いることで、配列データの被覆度を染色体特異的に高くすることが可能であり、配列データの解析において情報処理にかかる時間と労力を削減できる上、得られた解析データの考察、解釈を進める上でも効率的である。反復配列や染色体間の重複領域が多く見られるヒトゲノムの解析においては、有効な方法と思われる。染色体特化型リシーケンス法は、ヒトX染色体の解析結果が既に報告されており、染色体レベルでの個人間の相違を検出するシステムを構築することを考えており、これらの解析を系統的に進めている研究チームは国内外においてまだ存在しない。

日本人の標準ゲノム配列は、今後の個人医療を実現するためには必要不可欠な基盤情報である。また、これまでにワトソン博士を始めとして4人の個人ゲノム配列が報告されており、本研究で得られた解析基盤を今後の個人ゲノム解析に適用し遺伝的多様性を明らかにすることにより、国際貢献ができるとともに医療への応用を目指したゲノム解析が飛躍的に進むことが予想される。

<達成できなかったこと、予想外の困難、その理由>

哺乳類性染色体の比較ゲノム研究を進める上で、原始哺乳類とも言える有袋類の性染色体のゲノム情報は、祖先型染色体の推測を行う上で有用である。しかし、霊長類と有袋類では進化距離が離れているため、ゲノム構造の相同性が低く、Y染色体の比較解析を行う上で対応領域の同定に苦労していた。今回、オーストラリアの研究チームと共同でタマウラビーY染色体由来のBACクローンをアンカーポイントに、ゲノムライブラリーのスクリーニングを行い、Y染色体に位置する新規クローンの単離に成功した。それらのクローンのうち、遺伝子を含むクローンについては、クローンの高精度塩基配列決定を行い、ヒト、チンパンジーなど、

高精度配列データが報告されている生物種のY染色体の配列と比較解析を行うべく、解析を進めている。また、今回解析を行った、タマウラビーY染色体由来のクローンは、Y染色体特異的反復配列を含むものが多く、さらにfiber-FISH法によるゲノム構造解析の結果、複数種のY染色体特異的反復配列が存在し、それぞれが、数百kbの領域に渡り、タンデムに反復していることがわかってきた。

次世代シーケンサーを運用するためのプロトコルの最適化と得られたデータの配列解析に難航し、結果を検証するためのシステムの構築が十分にできなかった。

<今後の課題>

これまで同定されたDNA構造の違いが日本人の集団では見いだせない(疾患に関与していない)場合が多いため、独自に日本人ゲノム配列を決定する必要がある。そこで、HapMapプロジェクトで使用されたJPTパネル(日本人由来)の細胞株を用いて一塩基多型(SNP)および構造多型などを検出し、その結果をデータベース化するとともにハプログループ間で比較する。

また今後、日本人の個人ゲノム解析を進めて行くうえで、ヒトゲノム・遺伝子解析に関する倫理指針の改正や個人ゲノム解析研究の意義について国民のコンセンサスを得ることが必要である。また、情報管理体制の構築も早急に進めていく必要がある。

<成果公表リスト>

- 0806261537
Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutiérrez EL, Dubchak I, Garcia-Fernández J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin-I T, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS. The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453, 1064-1071 (2008)
- 0901150042
Kumagai H, Utsunomiya S, Nakamura S, Yamamoto R, Harada A, Kaji T, Hazama M, Ohashi T, Inami A, Ikegami T, Miyamoto K, Endo N, Yoshimi K, Toyoda A, Hattori M, Sakaki Y.: Large-scale microfabricated channel plates for high-throughput, fully automated DNA sequencing. *Electrophoresis*. 23, 4723-32 (2008)
- 0901140104
Hongoh Y, Sharma VK, Prakash T, Noda S, Toh H, Taylor TD, Kudo T, Sakaki Y, Toyoda A, Hattori M, Ohkuma M.: Genome of an endosymbiont coupling N2 fixation to cellulolysis within protist cells in termite gut. *Science* 322, 1108-9 (2008).
- 0805032052
Hongoh Y, Sharma VK, Prakash T, Noda S, Taylor TD, Kudo T, Sakaki Y, Toyoda A, Hattori M, Ohkuma M.: Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc Natl Acad Sci U S A*. 105, 5555-60 (2008)

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

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

1) 国立情報学研究所 2) 国立遺伝学研究所 3) 東京大学大学院新領域創成科学研究科 4) 九州大学大学院農学研究院
5) 理化学研究所 6) 東京大学大学院理学研究科 7) 東京大学大学院農業生命科学研究科 8) 名古屋大学大学院理学研究科

<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する方がふさわしい、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。

<2008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で下記の要領で実施したが、構成班員の移動と研究機関の組織改編、科学研究費の実施要領の変更に伴い、分担班員と連携班員に変更が生じている。また、昨年度の東京大学に引き続き、国立遺伝学研究所にもILLUMINA-GAIIゲノムアナライザを導入し、これを用いた支援を新たに開始している。

(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、メールリストを利用し全班員に通知した。

(3) 支援依頼は、本班代表者宛のメー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) 各班員への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で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だき、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。

<2008年度の成果>

昨年度までに27件の支援依頼を受け付けたが、今年度の支援依頼数は大幅に増加し、平成21年1月までに26件である。全領域において班員の研究の進捗状況の反映と思われる。各々の支援依頼については、支援委員会で審査後に比較ゲノム支援班、基盤ゲノム小原班、菅野班が連携し、申請班員と具体的な支援方法について協議の上で支援を行っている。これまでに受け付けた53件の支援依頼のうち52件については、何らかの形で支援を実施しているが、経費と人員の両面で支援能力の上限に近づきつつある。

シーケンシング支援

全ゲノムを対象：3

BAC・フォスミド・cDNAクローン全長配列決定：11

BAC・フォスミドライブラリ末端配列決定：1

cDNAライブラリエンド配列決定：6

PCR産物を用いたゲノムリシーケンシング：2

GAIIによる支援(ゲノム・メタゲノム)：6

GAIIによる支援(トランスクリプト)：2

ライブラリ・リソース構築支援

BAC・フォスミドライブラリ作成：4

完全長cDNAライブラリ作成：2

マイクロアレイ技術支援：1

不採択：1