

厚生労働科学研究費補助金

ヒトゲノム・再生医療等研究事業

サル完全長cDNAの配列決定とヒト遺伝子との比較解析および
配列情報に基づくcDNAアレイ作製と応用に関する研究

平成16年度 総括・分担研究報告書

主任研究者 橋本雄之

平成17(2005)年3月

平成16年度

厚生科学研究費補助金（ヒトゲノム・再生医療等研究事業）総括研究報告書

サル完全長 cDNA の配列決定とヒト遺伝子との比較解析および配列

情報に基づくcDNAアレイ作製と応用に関する研究

課題番号：H15-ゲノム-003

主任研究者：橋本 雄之 国立感染症研究所遺伝子資源室長

カニクイザルcDNA約7万クローンの5'端配列を元に、ヒト参照(RefSeq)遺伝子配列に対してホモロジー検索を行ったところ、45,000クローンが8,200種のヒト参照遺伝子に対応した。それらの全長配列を台湾アカデミアシニカのグループと共同で進め、脳由来約5,000、精巣由来約2,000の全長配列を決定し、ヒトとの配列比較解析を行った。その結果、脳で発現している遺伝子約2,000についてみると、他の組織の値と比べて、Ks(同義置換率)はほぼ6%と同じなのに対し、Ka(非同義置換率)は脳で1%と他組織に比べて1/2になっていることが分かった。さらに、ヒトではカニクイザル、チンパンジーに比べて、平均的にタンパク質の配列がより保存されていることも分かった。次に、ヒト相同遺伝子をもつカニクイザルcDNAについて、その5'UTRを比較すると、配列がアラインできないような領域が存在し、その領域の相同配列はヒトゲノム配列には存在するが、通常のヒト転写物にみられなかった。そうしたエクソン領域をスタートエクソンとするなどの変化型5'UTRが精巣cDNAでは27%の割合で存在したが、脳由来cDNAでは14%でそれほど高くなかった。ヒトやチンパンジーの精巣転写物にはそのような高率で変化型が見られないので、カニクイザル精巣特有のことと思われる。

チンパンジー全長cDNA配列について、ヒトとの配列比較を行ったところ、塩基置換率は5'UTRで0.99%、翻訳領域(CDS)で0.61%、3'UTRで0.90%、アミノ酸配列(aa)で0.55%、cDNA配列全体で0.72%であった。アミノ酸置換が起こった遺伝子は74%におよぶ。挿入・欠失(indel)に着目して解析した結果、ヒトとチンパンジー間で48%の遺伝子にindelが存在し、5'UTR、3'UTR、CDSの順で頻度が高いことがわかった。indelを合わせると、ヒトとチンパンジー間での配列の違いは塩基置換のみの値に対して全体に約2割増しの結果となった。

疾患遺伝子の機能と霊長類の進化との関係を知るため、アミノ酸置換と発症との関係が明白な1300個のヒト疾患関連遺伝子に相補性をもつ309個のカニクイザル遺伝子cDNAを全長配列決定遺伝子コレクションから選別した。このなかで、ヒトでは病因となるアミノ酸置換がカニクイザル遺伝子の正常配列となる例が9遺伝子で10カ所に認められ、これらは発症機構の解明に役立つものと思われる。

カニクイザル、チンパンジー、ヒトの遺伝子の転写開始部位をゲノムにマップし、プロモーター部位の同定を行い、チンパンジーについては400プロモーターを、カニクイザルについては300プロモーターを同定した。また、これらのプロモーターのクローニングを行った。

分担研究者

楠田 潤 国立感染症研究所 主任研究官

平井 百樹 東京女子医科大学

心臓血管研究所 教授

菅野 純夫 東京大学大学院

新領域創成科学研究科 教授

A. 研究目的

ヒトゲノム解析プロジェクトの進展で、ヒトゲノム全塩基配列が利用可能となり、ヒト完全長cDNAやESTの配列情報を基にヒト遺伝子2万以上が国際データベースに記載されている。しかし、ゲノムの遺伝子部分の確定とその機能解明のために、分離した完全長cDNAクローンとその配列情報を研究の材料とすることは依然として必要である。

本研究ではヒトに近いモデル動物としてのカニクイザル完全長cDNAコレクションに基づくヒトゲノム遺伝子領域解析をめざして、1) カニクイザル脳7領域、精巣および肝臓から、オリゴキャップ法により作製したライブラリー由来cDNAクローンをもとに、ヒト参照遺伝子(RefSeq)配列に相当するものを検索し、その全長塩基配列を決定し、ヒトとの配列比較解析を行う。チンパンジーは配列一致率が高すぎて意味のある差異の解析ができない例が多いので、カニクイザルを主とするが遺伝子によって適宜チンパンジーcDNAも比較に加えて、種特異的に出現(欠失)している遺伝子や種特異的進化を示す遺伝子の探索など、遺伝子進化の異同を解析し、ヒトの進化的成り立ちを遺伝子レベルで追究する。特に、ヒトの遺伝疾患として記載のあるもの約1600個についてはヒト参照遺伝子のcDNAクローンを新たに配列決定して、サルcDNAとの配列比較により進化的差異を解析し、種による遺伝子機能変化と病因との関係を探る。それらをもとに遺伝子機能探索のために疾病解析用完全長cDNAバンクの確立をはかり、疾病遺伝子cDNAの全長配列に疾病との関連データを付加したデータベース作りを行うとともに、cDNAクローンのバンクでの運用を可能にする。2) 上記で分離したカニクイザルcDNAの配列情報に基づき、2万種の遺伝子に対応するオリゴヌクレオチドマイクロアレイ作製を試み、脳部位での特定遺伝子発現の差異や肝臓での薬剤応答遺伝子

発現変動を解析することなどへの利用のためバンクからcDNAアレイで供給できるようにする。

これによってこれまで不明だったヒトの疾病に関連する遺伝子の同定から機能解明に向けた研究、また、既に関係の知れたものも含めて遺伝子の機能変化とその疾病の成因との関係を解析する研究、さらに、診断、治療に結び付く研究の発展に資することを目的とする。

B. 研究方法

1) サル脳各部の7組織(頭頂葉、側頭葉、前頭葉、後頭葉、脳幹、小脳皮質、延髄、)及び精巣、肝臓からオリゴキャップ法により作製した完全長cDNAに富むライブラリーDNAから分離し、部分塩基配列を決定したクローンについて、相同性検索によりヒト参照遺伝子(cDNA)配列に相当するものについて、その全長塩基配列を年3,000個決定(一部、台湾アカデミアシニカ分子生物学研究所との共同研究)し、カニクイザル参照遺伝子セットを作製する。蛋白コード領域を中心としてヒトとの配列比較を行い、遺伝子進化の異同を解析し、種の独自性との関連を追究する。特に、ヒト疾患遺伝子に対応する配列を検索し、対応するサルcDNA全長配列決定を行い、ヒトで病気を引き起こす遺伝子の変化と進化過程での変化の比較解析を進め、配列の差と病因との関係を探る。また、疾病遺伝子cDNAの全長配列に疾病との関連データを付加したデータベース作りを行うとともに、cDNAクローンを用いて機能解析ができるようにヒト疾患関連遺伝子cDNAも新たにそろえ、カニクイザルcDNAとセットで疾病解析用遺伝子バンクの確立に努める。(橋本、楠田、菅野)

2) チンパンジー組織が入手できたので、新たに精巣cDNAライブラリーを作製してcDNA数(約1万)を増やして、5'端配列

を決定する。そのデータをふまえて、上記の配列比較に適切なチンパンジーcDNAを加えることにより、スプライシング変異や挿入・欠失などの構造変異、塩基（アミノ酸）置換率、非同義置換数と同義置換数の比などを調べ、300以上の遺伝子について種間での配列保存度を比較することにより、種特異的に出現（欠失）している遺伝子や種特異的進化を示す遺伝子の探索を行う。また、チンパンジー大脳、小脳、腎臓、肝臓などの組織よりtotal RNAを抽出し、ヒトならびにカニクイザルのマイクロアレイを用いた解析のための試料とする。（平井、橋本）

3) 肝臓由来cDNAクローンを含めて、カニクイザルcDNAクローン9万個について、その配列情報をもとに、オリゴヌクレオチドを設計し、特異性が高く、感度のよいオリゴヌクレオチドマイクロアレイを試作する。最終的に、約2万種の遺伝子に対応するオリゴヌクレオチドマイクロアレイ作製を試み、脳部位での特定遺伝子発現の差異や肝臓での薬剤に応答した遺伝子発現変動の解析など、その利用法を探りバンクから供給できるようにする。（菅野、橋本）

（倫理面への配慮）

ヒトcDNAライブラリー原材料は医療機関に属する医師により、本人または遺族の書面による許諾を得て、採取されたもので、コード化され個人識別情報は研究者に示されない。さらに、作製に当たって複数の材料を使用するなどの手段で保護を行う形を取った。

カニクイザルは国立感染症研究所サル需給調整委員会及び動物実験委員会で審議されたもので、殺処分を行わざるをえなくなったものを霊長類センター研究員との共同研究という形で分与を受けた。チンパンジー組織は三和科学所有で自然死したものから、同所倫理委での議を経て分与を受けた。

C. 研究結果

1) これまでに分離した約7万クローンのうち、その5'端配列を元に、ヒト参照(RefSeq)遺伝子配列に対して、ホモロジー検索を行ったところ、45,000クローンがヒト参照遺伝子(RefSeq)の8,200種に対応した。それらの代表クローンの全長配列を台湾アカデミアシニカのグループと共同で進め、脳由来約5,000、精巣由来約2,200の全長配列を決定した。ヒトとの配列比較解析を行った。その結果、1)

全長配列のヒト遺伝子配列との比較解析の結果、脳で発現している遺伝子約2,000についてみると、他の組織の値と比べて、Ks(同義置換率)はほぼ6%と同じなのに対し、Ka(非同義置換率)は脳で1%と他組織に比べて1/2になっていることが分かった。脳では平均的に遺伝子の変化が少なくなっているとみられる。ヒトではカニクイザル、さらにチンパンジーに比べて、平均的にタンパク質の配列がより保存されていた。ヒト脳の高度の機能を維持するために他の遺伝子産物との相互作用の維持などの機能的制約があって、遺伝子(タンパク質配列)の進化速度が他に比較して遅くなっていると考えられる。

2) ヒト相同遺伝子をもつカニクイザルcDNAについて、その5'UTRを比較すると、配列がアラインできないような領域が存在した。その領域の相同配列はヒトゲノム配列には存在するが、通常のヒト転写物にみられない。そうしたエキソン領域をスタートエキソンとするなどの変化型5'UTRが精巣cDNAでは27%の割合で存在することが分かった。脳由来cDNAでは14%でそれほど高くなかった。ヒトやチンパンジーの精巣転写物にはそのような高率で変化型が見られないので、カニクイザル精巣特有のことと思われる。

2) 疾患遺伝子の機能と霊長類の進化

との関係を知るため、アミノ酸置換と発症との関係が明白な1300個のヒト疾患関連遺伝子に相補性をもつ309個のカニクイザル遺伝子 cDNA を全長配列決定遺伝子コレクションから選別した。このなかで、ヒトでは病因となるアミノ酸置換がカニクイザル遺伝子の正常配列となる例が9遺伝子で10カ所に認められたが、これらは発症機構の解明に役立つものと思われる。

3) チンパンジーの完全長cDNAライブラリーを作製し、約1万5千クローンから得られた5'端配列について、3,771種のヒト既知mRNAに対応する配列が得られた。それらのうち約210配列について全長配列を決定した。チンパンジー全長cDNA配列について、ヒトとの配列比較をおこなったところ、塩基置換率は5'非翻訳領域(5'UTR)で0.99%、翻訳領域(CDS)で0.61%、3'非翻訳領域(3'UTR)で0.90%、アミノ酸配列(aa)で0.55%、cDNA配列全体で0.72%であった。アミノ酸置換が起こった遺伝子は74%(65/88)におよぶことが確認できた。挿入・欠失(indel)に着目して解析した結果、ヒトとチンパンジー間で42%(42/88)の遺伝子にindelが存在し、5'UTR、3'UTR、CDSの順で頻度が高いことがわかった。興味深いことに、両UTRではヒト欠失(チンパンジー挿入)、逆にCDSではヒト挿入(チンパンジー欠失)の頻度が高い傾向が見られた(両UTRでは約2倍、CDSでは約3倍の差)。indelを合わせると、ヒトとチンパンジー間での配列の違いは塩基置換のみの値に対して全体に約2割増しの結果となった。

4) これまで得られた配列データを、チンパンジー、カニクイザル等のゲノム配列にマップし、チンパンジーでは約400の、カニクイザルでは約300の遺伝子について、転写開始部位についてマップすることができた。これらの転写開始部位を含む約1,000塩基のゲノム領域をチンパンジーおよびカ

ニクイザルのゲノムDNAからクローン化をおこない、その活性をヒト培養細胞HEK293で検討したところ、ほとんどすべてのものにプロモーター活性が認められた。

上記カニクイザルcDNA配列情報をもとに、オリゴヌクレオチドを設計し、ヒト参照遺伝子に対応する約2,000の遺伝子についてオリゴヌクレオチドマイクロアレイ作製を試み、脳部位での特定遺伝子発現の差異や肝臓での薬剤に応答した遺伝子発現変動の解析を可能とした。

D. 考察

カニクイザルはニホンザルと同属で、性格がより穏やかで扱いやすい為、対人薬の開発などに実験動物として用いられているが、網羅的遺伝子解析は行われていなかった。ゲノムからヒト化の過程を明らかにするには近縁の霊長類での比較解析が不可欠であり、特に、霊長類としての特徴が明確になる旧世界ザルでの解析が重要である。日本でチンパンジーでのゲノム配列解析がヒトとの比較で進められはじめたが、大規模cDNA解析は行われていない。一方、ヒトとマウスでcDNA多数が登録されているが、完全長cDNAクローンを整備して、機能を探ることは疾患の成立を研究するうえで有効であり、ゲノム上に遺伝子を確定するのに必須と世界的に注目されてきた。カニクイザルはマウスなどよりもはるかにヒトに近縁であり、各遺伝子構造はヒトとほぼ同じであると考えられている。脳は、体の中で最もバリエーションに富んだ遺伝子が発現しており、精巣では臓器特異的に発現して新規に分離されるcDNA種が多いと期待された。さらに、カニクイザルcDNAライブラリーを用いることにより、ヒトよりも新鮮なサンプルが入手でき、ヒトのcDNAライブラリーでは作製の過程で失われる可能性のある希少な発現の遺伝子を発見することが期待できる。

ヒト遺伝子配列と相同性のあるカニクイザルcDNAクローンについて、ORFがほぼ一致する精巣由来クローンについて、塩基置換率を求めたところ、精巣由来cDNAで非同義部位の塩基置換率値が高い、すなわち、進化の過程でアミノ酸変化を起こしてきた度合いが高いことが示唆された。これは5' UTR領域でのエクソン変化とあわせて、遺伝子発現変化への影響ということで興味あるデータである。

アミノ酸置換と発症との関係が明白な1300個のヒト疾患関連遺伝子に相補性をもつ309個のカニクイザル遺伝子cDNAを全長配列決定遺伝子コレクションから選別したなかで、ヒトでは病因となるアミノ酸置換がカニクイザル遺伝子の正常配列となる例が9遺伝子で10カ所に認められたが、これらはタンパクの進化と発症機構の解明に役立つものと思われる。

E. 結論

カニクイザル脳各部および精巣から完全長cDNAに富む7万クローンを分離・保存し、その5'端配列を元に、ヒト参照(RefSeq)遺伝子配列に対して、ホモロジー検索を行ったところ、45,000クローンがヒト参照遺伝子(RefSeq)の8,200種に対応した。それらの代表クローンの全長配列を台湾アカデミアシニカのグループと共同で進め、脳由来約5,000、精巣由来約2,200の全長配列を決定した。

疾患遺伝子の機能と霊長類の進化との関係を知るため、アミノ酸置換と発症との関係が明白な1300個のヒト疾患関連遺伝子に相補性をもつ309個のカニクイザル遺伝子cDNAを全長配列決定遺伝子コレクションから選別した。このなかで、ヒトでは病因となるアミノ酸置換がカニクイザル遺伝子の正常配列となる例が9遺伝子で10カ所に認められた。

チンパンジーの完全長cDNAライブラリ

一を作製し、約1.5万クローンから得られた5'端配列について、3,771種のヒト既知mRNAに対応する配列が得られた。それらのうち約210配列について全長配列を決定した。チンパンジー全長cDNA配列について、塩基置換率は5'非翻訳領域(5'UTR)で0.99%、翻訳領域(CDS)で0.61%、3'非翻訳領域(3'UTR)で0.90%、アミノ酸配列(aa)で0.55%、cDNA配列全体で0.72%であった。アミノ酸置換が起こった遺伝子は74%によぶ。挿入・欠失(indel)に着目して解析した結果、ヒトとチンパンジー間で48%の遺伝子にindelが存在し、5'UTR、3'UTR、CDSの順で頻度が高いことがわかった。indelを合わせると、ヒトとチンパンジー間での配列の違いは塩基置換のみの値に対して全体に約2割増しの結果となった。

これまで得られた配列データを、チンパンジー、カニクイザル等のゲノム配列にマップし、チンパンジーでは約400の、カニクイザルでは約300の遺伝子について、転写開始部位についてマップすることができた。これらの転写開始部位を含む約1000塩基のゲノム領域をチンパンジーおよびカニクイザルのゲノムDNAからクローン化をおこない、その活性をヒト培養細胞HEK293で検討したところ、ほとんどすべてのものにプロモーター活性が認められた。

F. 研究発表

1. 論文発表

1. Suzuki, Y., Yamashita, R., Sugano, S. and Nakai, K. DBTSS, DataBase of Transcriptional Start Sites: progress report 2004. *Nucleic Acids Res.* 32: D78-81, 2004.

2. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, Yura K, Miyazaki S, Ikeo K, Homma K, Kasprzyk A, Nishikawa T, Hirakawa M, Thierry-Mieg J, Thierry-Mieg D, Ashurst J, Jia L, Nakao M, Thomas MA, Mulder N, Karavidopoulou Y, Jin L, Kim S, Yasuda T, Lenhard B, Eveno E, Suzuki Y, Yamasaki C, Takeda JI, Gough C, Hilton P, Fujii Y, Sakai H, Tanaka S, Amid C, Bellgard M, Bonaldo Md M, Bono H, Bromberg SK, Brookes AJ, Bruford E, Carninci P, Chelala C, Couillault C, Souza SJ, Debily MA, Devignes MD, Dubchak I, Endo T, Estreicher A, Eyraas E, Fukami-Kobayashi K, R Gopinath G, Graudens E, Hahn Y, Han M, Han ZG, Hanada K, Hanaoka H, Harada E, Hashimoto K, Hinz U, Hirai M, Hishiki T, Hopkinson I, Imbeaud S, Inoko H, Kanapin A, Kaneko Y, Kasukawa T, Kelso J, Kersey P, Kikuno R, Kimura K, Korn B, Kuryshev V, Makalowska I, Makino T, Mano S, Mariage-Samson R, Mashima J, Matsuda H, Mewes HW, Minoshima S, Nagai K, Nagasaki H, Nagata N, Nigam R, Ogasawara O, Ohara O, Ohtsubo M, Okada N, Okido T, Oota S, Ota M, Ota T, Otsuki T, Piatier-Tonneau D, Poustka A, Ren SX, Saitou N, Sakai K, Sakamoto S, Sakate R, Schupp I, Servant F, Sherry S, Shiba R, Shimizu N, Shimoyama M, Simpson AJ,

Soares B, Steward C, Suwa M, Suzuki M, Takahashi A, Tamiya G, Tanaka H, Taylor T, Terwilliger JD, Unneberg P, Veeramachaneni V, Watanabe S, Wilming L, Yasuda N, Yoo HS, Stodolsky M, Makalowski W, Go M, Nakai K, Takagi T, Kanehisa M, Sakaki Y, Quackenbush J, Okazaki Y, Hayashizaki Y, Hide W, Chakraborty R, Nishikawa K, Sugawara H, Tateno Y, Chen Z, Oishi M, Tonellato P, Apweiler R, Okubo K, Wagner L, Wiemann S, Strausberg RL, Isogai T, Auffray C, Nomura N, Gojobori T, Sugano S. :

Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones.

PLoS Biology 2;6: 0856-0875 , 2004.

3. Suzuki Y, Yamashita R, Shiota M, Sakakibara Y, Chiba J, Mizushima-Sugano J, Nakai K, Sugano S. Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions. *Genome Res.* 14:1711-1718, 2004.

4. Bajic VB, Tan SL, Suzuki Y, Sugano S. Promoter prediction analysis on the whole human genome. *Nat Biotechnol.* 22: 1467-1473, 2004

5. Suzuki Y, Yamashita R, Shiota M, Sakakibara Y, Chiba J, Mizushima-Sugano J, Kel AE, Arakawa T, Carninci P, Kawai J, Hayashizaki Y, Takagi T, Nakai K, Sugano S. Large-scale collection and characterization of promoters of human and mouse genes. *In Silico Biol.* 4(3):0036 [Epub ahead of print] 2004

2. 学会発表

1) 坂手龍一, 今西 規, 平井百樹, 橋本雄之, 五條堀 孝: ヒトとチンパンジーのcDNA配列の比較解析から探るヒト遺伝子の進化
日本遺伝学第76回 大会、吹田、2004年9月

2) 橋本雄之、長田直樹、平田誠、田沼玲子、楠田潤: カニクイザル精巢cDNAとヒト相同遺伝子との配列比較—27%の5' UTRに変化
第27回日本分子生物学会年会、神戸、2004年12月

3) 楠田潤、平田誠、田沼玲子、長田直樹、平井百樹、橋本雄之: 神経疾患遺伝子の霊長類ホモログに関する研究
第27回日本分子生物学会年会、神戸、2004年12月

4) Y. Suto , M.H. Park, D.H. Seo, M. Hammond, E. Smart, R. Matsuoka, M. Hirai, T. Juji and Y. Ishikawa : Fiber fluorescence in situ hybridization (FISH) analysis of Rhesus (Rh) blood group locus. Human Genome Variation Society Scientific & Annual General Meeting 2004 (26 October, Toronto, Canada)

G. 知的所有権の取得状況

取得なし

平成 16 年度 厚生労働科学研究費補助金（ヒトゲノム・再生医療等研究事業）
分担研究報告書

カニクイザル完全長cDNAの配列決定とヒトオルソログとの比較解析

橋本雄之

国立感染症研究所 遺伝子資源室

これまでに分離した約7万クローンについて、その5'端配列をもとに、ヒト参照(RefSeq) 遺伝子配列に対して、ホモロジー検索を行ったところ、4,5000 クローンがヒト参照遺伝子(RefSeq) の 8,200 種に対応した。それらの代表クローンの全長配列を決めることを台湾アカデミアシニカのグループと共同で進め、脳由来約 5,000、精巣由来約 2,000 の全長配列を決定し、公共データベースに登録した。

全長配列のヒト遺伝子配列との比較解析の結果、脳で発現している遺伝子約2,000についてみると、他の組織の値と比べて、Ks (同義置換率) はほぼ6%と同じなのに対し、Ka (非同義置換率) は脳で1%と他組織に比べて1/2になっていることが分かった。また、ヒトではカニクイザル、さらにチンパンジーに比べて、平均的にタンパク質の配列がより保存されていることも分かった。ヒト脳の高度の機能を維持するために他の遺伝子産物との相互作用の維持の必要性などの機能的制約があつて、遺伝子の進化速度が他に比較して遅くなっていると考えられる。

ヒト相同遺伝子をもつカニクイザルcDNAについて、ORF がほぼ一致する精巣由来クローン 624 ペア、脳由来 307 ペアについて、その5' UTRを比較すると、配列がアラインできないような領域が存在し、その領域の配列はヒトゲノム配列には存在するが、通常のヒト転写物にみられない。そうしたエクソン領域をスタートエクソンとするなどの変化型5' UTRが精巣cDNAでは27%の割合で存在するが、脳由来cDNAでは14%でそれほど高くなかった。ヒトやチンパンジーの精巣転写物にはそのような高率で変化型が見られないので、カニクイザル精巣特有のことと思われた。

A. 研究目的

ヒトゲノム全体のDNA配列決定プロジェクトとその一環として発現遺伝子部分のクローン化が大規模に進められているが、遺伝子部分の確定とその機能解明は、全体としては残された課題となっている。

1) カニクイザル脳 7領域、精巣および肝臓から、オリゴキャップ法により作製したライブラリー由来cDNAクローン約9万個をもとに、その3万個のクラスターのうち、ヒト参照遺伝子(RefSeq)配列に相当するものを検索し、有用なものから年3,000個の全長塩基配列を決定し、ヒトとの配列比較解

析を行う。チンパンジーは配列一致率が高すぎて意味のある差異の解析ができない例が多いので、カニクイザルを主とするが遺伝子によって適宜チンパンジーcDNAも比較に加えて、種特異的に出現(欠失)している遺伝子や種特異的進化を示す遺伝子の探索など、遺伝子進化の異同を解析し、ヒトの進化的成り立ちを遺伝子レベルで追究する。さらに、ヒト疾病関連遺伝子の同定から機能解明、その疾病の成因解明そして診断、治療に結び付く研究の発展に資する。

B. 研究方法

1) 自動DNAシーケンサーを用いて、クローンDNAの部分塩基配列を決定し、インサートされたDNAの5'端から部分塩基配列決定を行った。多数のcDNAクローンについてコンピュータープログラムによりDNAデータベース上の配列と相同性検索を行い、既知のものに合致するものについて、その5'領域を完全に含んでいるかどうか判定した。

3) ヒト既知配列と相同性のあるクローンのインサートについて、全長配列をプライマーウオーキング法を主として決定した。それらの全長配列について、米国UCSCヒトゲノム配列データベースのBlat programによる相同性検索を行った。配列比較はClustalWプログラムを用いて並列配列を作り、Liの方法で塩基置換率Ka, Ks値を求めた。K5'UTRはKimura's 2-parameter法によって求めた。

C. 研究結果

サル脳各部の7組織(頭頂葉、側頭葉、前頭葉、後頭葉、脳幹、小脳皮質、延髄)及び精巣、肝臓からオリゴキャップ法により作製した完全長cDNAに富むライブラリーDNAから分離し、部分塩基配列を決定したクローンについて、相同性検索によりヒト参照遺伝子(cDNA)配列に相当するものについて、その全長配列を決定し、カニクイザル参照遺伝子セットを作製することを目標とした。これまでに7,000クローンについて決定(一部、台湾アカデミアシニカ分子生物学研究所との共同研究)し、公共データベースに登録した。

全長配列のヒト遺伝子配列との比較解析の結果脳で発現している遺伝子約2,000について塩基置換率を調べると、他の組織の値と比べて、表1に示したように、Ks(同義置換率)はほぼ6%と同じなのに対し、

Ka(非同義置換率)は脳で1%と他組織に比べて1/2になっていることが分かった。脳では平均的に遺伝子の変化が少なくなっているとみられる。また、ヒトではカニクイザル、さらにチンパンジーに比べて、平均的にタンパク質の配列がより保存されていることも分かった。ヒト脳の高度の機能を維持するために他の遺伝子産物との相互作用の維持の必要性などの機能的制約があって、遺伝子(タンパク質配列)の進化速度が他に比較して遅くなっていると考えられる。

ヒト相同遺伝子をもつカニクイザルcDNAについて、その5'UTRを比較すると、配列がアラインできないような領域が存在した。その領域の相同配列はヒトゲノム配列には存在するが、通常のヒト転写物にみられない。そうしたエキソン領域をスタートエキソンとするなどの変化型5'UTRが精巣cDNAでは27%の割合で存在することが分かった。脳由来cDNAでは14%でそれほど高くなかった。ヒトやチンパンジーの精巣転写物にはそのような高率で変化型が見られないので、カニクイザル精巣特有のことと思われる。カニクイザル特有の精巣でのタンパク質発現制御にとって必要になったと考えられる。

こうしたヒト遺伝子配列と相同性のあるクローンのうち、ORFがほぼ一致するもの、精巣由来624、脳由来307ペアについて、塩基置換率を求めたところ(表2参照)、同義置換座位及び非同義置換座位における塩基置換率(Ks及びKa)は脳で有意に低く、同じく5'UTRでの塩基置換率も低くなっていた。従って、カニクイザル精巣では一定の遺伝子群において変異が維持される必要が生じ、上記のエキソン変化と併せて、そのような変化が残る淘汰が働いたものと考えられる。

このカニクイザル特有の5'エキソンが実

際、ヒト転写物に存在しないかどうか、ランダムに18個を選び、特異的プライマーを作製して、ヒトおよびサニクイザル組織トータルRNAについてRT-PCRを行った。その結果、1/3についてはヒト組織では発現がみられず、カニクイザル特有に使用されるエクソンであることが示された。また、2/3のヒトで発現がみられたものでも、これまでのcDNAデータベースにはその部分を含んだcDNAは登録されてないわけで、ヒトcDNA分離もまだ不十分であることが示された。

D. 考察

非同義置換は精巣由来配列で高く、精巣で発現している遺伝子には進化過程で変化したものが多いことが示唆される。同じく5'UTRでのエクソン利用変化も高くなっていることと併せてみると、カニクイザル精巣では一定の遺伝子群において変異が維持される必要が生じ、そのような変化が残る淘汰が働いたものと考えられる。精巣では他組織で発現していない遺伝子が多く、cDNAとして分離されることと併せて大変興味深い。

E. 結論

これまでに分離した約7万クローンのうち、その5'端配列を元に、ヒト参照(RefSeq)遺伝子配列に対して、ホモロジー検索を行ったところ、45,000クローンがヒト参照遺伝子(RefSeq)の8,200種に対応した。それらの代表クローンの全長配列を決めることを台湾アカデミアシニカのグループと共同で進め、脳由来約5,000、精巣由来約2,200の全長配列を決定した。こうしたヒト遺伝子配列と相同性のあるクローンについて塩基置換率など、比較解析した。脳で発現している遺伝子約2,000

についてみると、他の組織の値と比べて、Ks(同義置換率)は同じなのに対し、Ka(非同義置換率)は脳では他組織に比べて1/2になっていることが分かった。また、ヒトではカニクイザル、さらにチンパンジーに比べて、平均的にタンパク質の配列がより保存されていることも分かった。ヒト相同遺伝子をもつカニクイザルcDNAについて、ORFがほぼ一致する精巣由来、脳由来ペアについて、その5'UTRを比較すると、配列がアラインできないような領域が存在し、その領域の相同配列はヒトゲノム配列には存在するが、通常のヒト転写物にみられない。そうしたエクソン領域をスタートエクソンとするなどの変化型5'UTRが精巣cDNAでは27%の割合で存在するが、脳由来cDNAでは14%でそれほど高くなかった。ヒトやチンパンジーの精巣転写物にはそのような高率で変化型が見られないので、カニクイザル精巣特有のことと思われた。

F. 研究発表

1. 論文発表

1) Imanishi, T. et al. (Hashimoto, K, Hirai, M. and Sugano, S.)

Integrative Annotation of 21,037 Human Genes Validated by Full-Length cDNA Clones

PLoS Biology 2;6: 0856-0875, 2004.

2. 学会発表

1) 坂手龍一, 今西規, 平井百樹, 橋本雄之, 五條堀孝: ヒトとチンパンジーのcDNA配列の比較解析から探るヒト遺伝子の進化

日本遺伝学第76回大会、吹田、2004年9月

2) 橋本雄之、長田直樹、平田誠、田沼玲子、楠田潤：カニクイザル精巣 cDNA とヒト相同遺伝子との配列比較— 27%の 5' UTR に変化
 第 27 回日本分子生物学会年会、神戸、2004 年 12 月

3) 楠田潤、平田誠、田沼玲子、長田直樹、平井百樹、橋本雄之：神経疾患遺伝子の霊長類ホモログに関する研究
 第 27 回日本分子生物学会年会、神戸、2004 年 12 月

表 1. ヒト・カニクイザル相同遺伝子 配列間での塩基置換率

ヒト/ カニクイザル	Ka (非同義置換率)	Ks (同義置換率)	Ka/Ks
brain	0.010	0.057	0.175
genome-wide	0.019	0.059	0.322

表 2.

cDNA source	N	Ka	Ks	K5	Ka/Ks	K5/Ks
Brain	307	0.0083	0.0581	0.0424	0.1429	0.7298
Testis	624	0.0146	0.0572	0.0482	0.2552	0.8427

平成16年度 厚生労働科学研究費補助金（ヒトゲノム・再生医療等研究事業）
分担研究報告書

カニクイザルとヒト疾病関連遺伝子比較解析

楠 田 潤 国立感染症研究所 遺伝子資源室

疾患遺伝子において疾病に関連したアミノ酸置換の進化的起源と哺乳類ホモログ間の分布を知るため、アミノ酸置換と発症との関係が明白な1,300個のヒト疾患関連遺伝子に相補性をもつ約300個のカニクイザル遺伝子を当研究室収蔵の完全長 cDNA 配列コレクションから選別した。このなかで、ヒトでは発症を促進もしくは抑制するアミノ酸置換がカニクイザル遺伝子の正常配列と一致する例が9遺伝子中に10箇所認められた。機能的には発症を促進するものが8置換で、抑制するものが2置換であった。これらの遺伝子の哺乳類ホモログをデータベース検索により収集し、変異アミノ酸を比較解析したところカニクイザルだけが疾患関連の変異を持つもの(6置換)、さらにヒトを除いてほとんどの哺乳類で変異型の配列をもつもの(3置換)、さらにカニクイザルより下等な動物では変異型をもつがそれより高等な種では正常型を示すもの(1置換)に分類された。特にヒト疾患の原因となる8置換はヒト以外の生物には何ら影響を与えないわけで、これらの置換を詳しく解析すれば疾患の発症機構の解明や治療薬の開発に役立つものと思われる。

A. 研究目的

ヒトゲノムには約25,000個の遺伝子が同定されているが、その変異が疾患の発症につながるような遺伝子を疾患遺伝子と呼んでいる。現在、Human Gene Mutation Database (HGMD) には1516個の疾患遺伝子が収録されており、疾患の原因となる39,415変異が登録されている。そのうち26,839個は nonsense/missense 変異、スプライシング、転写、翻訳に関する単一塩基置換であり、残りの12,576個は挿入/欠失変異であった。特に nonsense 変異によってもたらされるアミノ酸置換の中には他の霊長類疾患遺伝子ホモログの正常配列と一致するものがあり、このような配列はヒトでは疾患を引き起

こすのに、非ヒト霊長類では生理的になんの影響を与えないことを意味している。ヒトにとって病気を起こす変異が非ヒト霊長類正常遺伝子配列にどの程度存在し、分布するのか、また進化的にどのように出現し、消失していったかを明らかにするため、約8,000個のカニクイザル完全長 cDNA 配列の網羅的解読により得られた配列情報から、病的置換と一致する配列をコンピューターで検索したところ、9個のカニクイザル疾患遺伝子に10個のアミノ酸置換が見つかった。さらに同定されたアミノ酸置換の進化的消長をみるため、Database 上の哺乳類のホモログ遺伝子についてアラインメントを行なった。

B. 研究方法

Human Gene Mutation

Database (HGMD) に収録されている 1516 個の疾患遺伝子から nonsense 変異によるアミノ酸置換が病因として記載されている 1,300 遺伝子について遺伝子の CDS 配列と疾患に関するアミノ酸置換情報を Manual で download した。チンパンジーの疾患遺伝子で cDNA 配列が得られないものについてはヒト cDNA 配列を UCSC 検索サイト上のチンパンジーゲノム配列にアラインメントして求めた。

C. 研究結果

これまで約 8,000 個のカニクイザル完全長 cDNA クローンについて挿入配列の全長を解読し、収録してきたが、その中から 1,300 個のヒト疾患遺伝子に対応する、300 個のカニクイザルオルソログ遺伝子を選別した。これらの遺伝子を機能別に分類してみると代謝系 (27%)、神経系 (26%)、内分泌系 (6%)、血球系 (6%) などが主な構成遺伝子であった。次にヒトに較べてカニクイザルで異なるアミノ酸配列がヒトで疾患に関連して置換したアミノ酸と一致するものを探したところ、9 個の遺伝子で 10 箇所が一致した (表 1)。ここで検索したのは 300 遺伝子であり、疾患遺伝子総数 1,300 の 23% にあたり、逆算するとカニクイザル疾患遺伝子全体ではおおよそ 39 個の遺伝子で 43 箇所のアミノ酸が病的アミノ酸と一致するものと予想される。

疾病に関係するアミノ酸配列を持つカニクイザル疾患遺伝子をみてみると高血圧の一要因となる AGT (angiotensinogen)、コレステロール値の上昇を引き起こす CETP (cholesterol ester transfer protein)、血液中の遊離脂肪酸を増加させる GPD2 (glycerol-3-phosphate dehydrogenase 2) などで、これらのアミノ酸の変化が直接重篤な疾患を引き起こすものではなく、体質の維持に不可欠で変化し得るアミノ酸の両者が多型としてヒト集団中に存在し、人種によって、両アミノ酸の比率が異なることが知られている。一方、良性の天疱瘡をつくる ATP2C1 (ATPase, Ca²⁺-sequestering)、ループス腎炎を引き起こす FCGR2A (Fc fragment of IgG)、白質脳症の原因遺伝子である EIF2B2 (eukaryotic initiation factor-2B)、ガラクトースエピメラーゼ欠損症の原因である GALE (galactose epimerase)、緑内障の要因である OPTN (optineurin)、などはアミノ酸の置換が直接発症につながるものである。さらにこれらはアミノ酸の置換によって誘起される機能的変化によって次の二つのグループに分けられる。一つはアミノ酸置換が発症を促進し、症状を悪化させるように働くもので AGT, ATP2C1, CETP, EIF2B2, FCGR2A (Q127K), GALE, GPD2, OPTN などが含まれる。もう一つのグループは変異型配列が発症を抑制したり症状を軽減するもので FCGR2A (R131H) や TTR などがこれに分類される。

またアミノ酸の置換が進化的にどのように出現し、哺乳類ゲノムに分布し、またどのように消失したかを調べるため、9個の疾患遺伝子の哺乳類ホモログをデータベース検索により収集し、アラインメントを行った。そして置換アミノ酸に相当する各哺乳類遺伝子のアミノ酸を表2にまとめた。これらは疾病関連置換がカニクイザルでだけ見られる遺伝子(ATP2C1, CETP, EIF2B2, FCGR2A(Q127K), GALE)とヒトを除いたほとんどの種で見られる遺伝子(GPD2, AGT, FCGR2)、カニクイザルまでは病的アミノ酸をもつがそれより高等な霊長類では正常型を示す遺伝子(TTR)に分類される。ただしAGT, CETP, FCGR2ではマウスホモログの相補性が低いので比較の対象からはずした。

D. 考察

昨年までは神経疾患遺伝子に対象をしぼって解析を進めてきたが、該当するカニクイザル配列が少なく、結果の解釈が難しかったので本年度は疾病関連遺伝子全体に範囲を広げ、疾病に係る変異を比較解析して、アミノ酸置換の進化的起源や変異によってもたらされる機能の変化を系統発生的に考察することを目指した。今回、カニクイザル疾患遺伝子の約23%について疾患関連のアミノ酸置換を検索し、9遺伝子中に10個の置換を同定したが、この頻度を基にすればカニクイザル遺伝子全体でも約39遺伝子43アミノ酸置換が出現する計算になり、本研究で得られた9遺伝子、10アミノ酸置換は貴重な一致例といえる。ヒトゲノム中にみられ SNP (single nucleotide

polymorphism)が他の非ヒト霊長類ゲノム中のSNPと一致する確率は極めて低いといわれており、疾病関連置換もSNPの一種だと考えればカニクイザルで疾病関連アミノ酸置換と一致した9遺伝子(10置換)というのは希で、機能的に制約を受けているのかもしれない。

本研究ではアミノ酸置換が機能の低下を起こす7遺伝子と機能の低下を防止するような2遺伝子を選別したが、後者は家計分析の過程でアミノ酸置換が機能の低下を起こす変異との二重変異として同定されたものである。このようなアミノ酸置換は発症の抑制や症状を軽減する働きがあり、創薬のターゲットとなるであろう。

病因関連アミノ酸置換の進化的出現の様子をみるため各種のホモログ配列をアラインメントしたが、ATP2C1, EIF2B2, GALE, CETP, FCGR2ではカニクイザルでのみ置換がみられ、カニクイザルがヒトを含む系統と分岐した後、カニクイザルの系統でのみ置換が起きたことを示している。そして、ヒトでの病的変異は回帰変異の一種と捉えることができる。一方、GPD2, AGT, FCGR2Aではホモロジーの低いマウスを除けば、ヒト以外の哺乳類ではすべて異常型のアミノ酸配列を示し、ヒトだけが正常型アミノ酸をもつ。すなわちチンパンジーからヒトに進化する過程で変異が起き、そのアミノ酸置換がヒトの生理に都合のよい機能を与えたことが推測される。もしこの仮説が当たっているならこれらはヒト化に貢献したアミノ酸の変異である可能性が高い。同様にTTRではカニクイザル以下の種

で変異型アミノ酸をもつが、オランウタンより高等な種では正常型となっている。すなはちその正常型アミノ酸が Great Ape にとって有利な機能を与えたのか、また異常型アミノ酸によってもたらされる機能が Great Ape にはもはや必要ではなくなったのかのいずれかであろう。

変異アミノ酸の機能に及ぼす影響がわかっているものについてみると大きく変化したのは GPD2 で AGT では中度の上昇がみられた。また GALE および CETP ではわずかな活性の低下が観察されただけであった。すなはち GALE と CETP における疾病関連置換はごくわずかな活性の低下をもたらすだけで、この変化はカニクイザルの生理にとって許容範囲内にあり、正常はな状態を損なうことはないが、高度に発達したヒトの生理にとっては支障をきたすものなのかもしれない。

E. 結論

アミノ酸置換と発症との関係が明白な 1,300 個のヒト疾患関連遺伝子に相補性をもつ約 300 個のカニクイザル遺伝子を当研究室収蔵の約 8,000 個の完全長 cDNA 配列コレクションから選別した。

ヒトでは発症を促進もしくは抑制するアミノ酸置換がカニクイザル遺伝子の正常配列と一致する例が 9 遺伝子中に 10 箇所認められた。

機能的には発症を促進するものが 8 置換で、抑制するものが 2 置換であった。

疾患に関係する変異アミノ酸を哺乳類ホモログ間で比較解析したところカ

ニクイザルだけが疾患関連の変異を持つもの(6 置換)、さらにヒトを除いてほとんどの哺乳類で変異型の配列をもつもの(3 置換)などがあった。後者は非ヒト霊長類のヒト化に貢献した可能性が示唆される。

疾患に関連したアミノ酸置換はそのアミノ酸が確実に機能直結していることを示すものでこれらの進化的起源と哺乳類ホモログにおける分布を明らかにすることは創薬の一つの基盤になるものと思われる。

F. 研究発表

学会発表

1. 楠田潤、平田誠、田沼玲子、長田直樹、平井百樹、橋本雄之：神経疾患遺伝子の霊長類ホモログに関する研究 第 27 回日本分子生物学会年会、神戸、2004年12月

カニクワイザル疾病遺伝子中にみられる疾病関連アミノ酸置換

アミノ酸置換	発症を促進し症状を悪化させる												発症を防止し症状を軽減する			
	GPD2	AGT	ATP2C	EIF2B2	GALE	CE1P	FCGR2A	OPTN	FCGR2A	TTR	FCGR2A	TTR	FCGR2A	TTR		
Human normal	His 100%	Met 100%	Ile 100%	Glu 100%	Gly 100%	Ile 100%	Gln 100%	Arg 100%	Arg 100%	Arg 100%	Arg 100%	Arg 100%	Arg 100%			
Pathological mutation	Arg 264	Thr 235(268)	Val 580	Gly 213	Glu 319	Val 405(422)	Lys 127(161)	Gln 545	His (131)165	His 104(124)	His 104(124)	His 104(124)	His 104(124)			
Animo acid position	FFA conc	AGT conc	ND	ND	CE binding	activity	ND	ND	ND	ND	ND	ND	ND			
Altered function	400%	115-140%			sightly lower	93%										
Chimpanzee	Arg 100%	Thr 99%	Ile 99%	Glu 99%	Gly 99%	Ile 99%	Gln 96%	Arg 97%	His 99%	Arg 97%	Arg 98%	Arg 98%				
Gorilla		Thr 99%														
Orangutan					Gly	98%		Arg		Arg	Arg	Arg	97%			
Cynomolgus monkey	Arg 99%	Thr 82%	Val	Gly 99%	Glu 93%	Val 93%	Lys 96%	Gln 89%	His 91%	His 89%	His 89%	His 89%	93%			
Marmoset		Thr 86%														
Bovine			Ile 98%													
Swine																
Mouse	Arg 93%(Glu)	59%	Ile 59%	Glu 96%	Gly 94%	(Lys) 80%	(Val) 24%	Gly 55%	His 79%	His 55%	His 55%	His 55%	76%			

カニクイザル疾病遺伝子配列に局在する疾病関連アミノ酸置換

発症を促進し症状を悪化させる									
AGT	angiotensinogen	NM_000029	106150	QmoA-10278	268(M235T)	CM920010	Hypertension, association with 高血圧		
ATP2C1	ATPase, Ca(2+)-sequestering	NM_014382	604384	AF181120	I580V	CM020634	Hailey-Hailey disease 良性天疱瘡		
CETP	cholesteryl ester transfer protein	NM_000078	118470	QtsA-14108	422(I405V)	CM994290	Higher HDL cholesterol level, association with 高HDLコレステロール症		
EIF2B2	eukaryotic initiation factor-2B	NM_014239	606454	NM_014239	E213G	CM013528	Leukoencephalopathy with vanishing white matter 白質脳症		
FCGR2A (CD32A)	Fc fragment of IgG	NM_021642	146790	NM_021642	161(Q127K)	CM980739	Immunoglobulin binding variant Lupus nephritis, protection against, association 腎		
GALE	galactose epimerase	NM_000153	606953	QtsA-19135	G319E	CM980813	Galactose epimerase deficiency ガラクトースエピメラーゼ欠損症		
GPD2	glycerol-3-phosphate dehydrogenase 2	NM_000408	138430	QtsA-16270	H264R	CM012769	Increased plasma FFA and glycerol levels		
OPTN	optineurin	NM_021980	602432	AB063036	R545Q	CM020164	Glaucoma 1, open angle 緑内障		
発症を防止し症状を軽減する									
FCGR2A (CD32A)	Fc fragment of IgG	NM_021642	146790	NM_021642	165(R131H)	CM960640	Immunoglobulin binding variant Lupus nephritis, protection against, association 腎		
TTR	transthyretin	NM_000371	176300	QccE-14396	124(R104H)	CM992976	Amyloidotic polyneuropathy アミロイド沈着症		

研究要旨

チンパンジーの完全長 cDNA 約 1 万 5 千クローンの 5' 端解析を進め、約 210 クローンについては全長塩基配列を決定した。さらにこれらの配列のうち、チンパンジー 88 遺伝子を選び、その全長配列とそれに対応するヒト mRNA 配列 (RefSeq) とを、公開されているヒトゲノム配列上にマッピングし、exon-intron 構造および exon 配列の種間比較をおこなった。塩基置換を指標にした種間差は、5' UTR、CDS、3' UTR においてそれぞれ 0.99%、0.61%、0.90% であった。アミノ酸置換率は 0.55% で、65 遺伝子 (74%) に置換がみられた。また、塩基挿入/欠失 (indel) を算入すると、配列の種間差は、5' UTR、CDS、3' UTR、アミノ酸の順に、1.14%、0.72%、1.11%、0.66% となった。それぞれ、ほぼ 2 割の種間差の増加をみた。領域ごとにみると、5' UTR ではチンパンジー挿入・ヒト欠失型が多くが見られた。CDS では逆で、3 塩基単位のヒト挿入・チンパンジー欠失型が多く、チンパンジーでのアミノ酸欠失が多く認められるという結果になった。このように indel を考慮に入れた解析により、ヒトとチンパンジーの遺伝子機能ならびに発現制御の違いに関連する可能性のある情報が得られた。

A. 研究目的

2003 年春にヒトゲノム解読が完了して以後、研究の流れはプロテオーム解析の方向へと向かってきた。しかし一方で、ゲノム構成の解析も新たな展開をみて、ヒトゲノムの多様性に関する研究成果が相次いで報告されている。たとえば、ゲノムの重複領域における segmental duplication の詳細な解析により、ゲノム上の個人差が予測以上に大きいことがわかってきた。さらに、そのような重複領域における SNP の密度からみても、人種差ともいえる集団差が確認されている。これらの差は、薬剤代謝に関わる表現型の集団差などに関連して、オーダーメイド医療とも結びつく可能性を示唆するものである。また、HIV 感染性やエイズ発症には、ケモカインの一種 CCL3L1 遺伝子のコピー数の個人差が関連するなどの研究も出てきた。こうした表現型にみられる集団差の成り立ちを知る上で、ヒトゲノムの祖先型を最も近い過去まで共有するチンパンジー・ゲノムをアウトグループとして用いた考察がなされるようになってきている。上述の CCL3L1 遺伝子のコピー数はチンパンジーでは大き

く、これがチンパンジーの AIDS 発症が生じにくいことの一因と考えられる。ヒトのゲノムの多様性に目を向ける際に、霊長類、特にチンパンジーの遺伝子・ゲノムデータは重要な情報をもたらすと考えられる。従来のゲノム比較においては、塩基置換を主たる指標として解析が行われてきた。しかし、挿入/欠失 (indel) に関しては、まだ十分なデータが集積されていない。そこで、本研究では、チンパンジー遺伝子の全長配列を読み、indel を考慮に入れた比較解析を行うこととした。

B. 研究方法

(1) ヒトの遺伝子配列のデータベース (RefSeq や LocusLink) を用いて、本研究で得たチンパンジーの塩基配列と相同な配列を持つヒトの遺伝子データを収集した。こうして、各相同遺伝子に関するアノテーション情報を参照して種間比較を行なった。具体的には、本研究では BLAST プログラムを用い、ヒト mRNA、EST、genome の各配列データベースに対して相同性検索を行って配列データを分類した。そのうえで、チンパンジー全長配列を読むべき遺伝

子を選択した。その全長配列データならびに対応するRefSeqをヒトゲノム(NCBI b35, UCSC hg16)へマッピングし、exon-intron構造およびexon配列の比較をおこなった。両種間で対応する配列間では塩基置換率を求めた。また、indelの出現領域を分類し、出現率に関しては挿入あるいは欠失の塩基の長さに関わり無く1 indelをもって1 eventとして算定した。

(2) ヒトのSNPに関する比較ゲノム学的解析のモデルとして、 β 3アドレナリン受容体遺伝子(ADRB3)の第1 exonのSNPに注目した。190位のTがCに置換したalleleが日本人で15%以上存在する。これにより遺伝子産物のN-末端から64番目のアミノ酸がトリプトファンからアルギニンへ置換(Trp64Arg)する。この変異は、はじめ非インシュリン依存性糖尿病(NIIDM)の罹患と関連して発見された。交感神経から分泌されるアドレナリンは、体脂肪の蓄積を抑え、脂肪を燃焼させる働きがある。(Trp64Arg)では、ベータ3アドレナリン受容体の機能が低くなる。データベース上、新世界ザル、アカゲザル、チンパンジー、オランウータンを含めた哺乳動物はすべて置換型であった。すなわちヒトの非置換型がむしろ例外的であることになる。そこでチンパンジー20個体についてゲノムDNAをもちいてexon1の塩基配列を解析した。

C. 研究結果

(1) 全長cDNA配列の種間比較解析:

約210のチンパンジー全長cDNAの配列を得た(表1)。これらのうち88のチンパンジー全長配列と、対応するヒトRefSeqをヒトゲノム(NCBI b35, UCSC hg16)へマッピングし、exon-intron構造およびexon配列の比較をおこなった。その結果、チンパンジーに特異的と推測される選択的スプライシングが14遺伝子に計20例認められた。また、チンパンジー3' UTRにAlu挿入が2遺伝子、50塩基以上の単位長を持つ反復配列の繰り返し数の違いによる大規模な挿入・欠失(indel)が2遺伝子に見つかった。また、転写開始点(5')は転写終了

点(3')よりも種間での保存性が低く、開始点はチンパンジーの方がヒトより上流に位置する傾向が見られた。Exonの塩基置換率は平均で0.72%であり、5' UTR、CDS、3' UTR、アミノ酸の順に、0.99%、0.61%、0.90%、0.55%であった。65遺伝子(74%)にアミノ酸置換がみられた。42遺伝子(48%)について60カ所、合計226bpのindelが見つかり、うち43indel(72%)は3' UTRに存在した。CDSは10indel(20%)にとどまるが、合計長は111bp(50%)に達した。Indelを加えると配列の種間多様性は、exon全体で0.86%、5' UTR、CDS、3' UTR、アミノ酸の順に、1.14%、0.72%、1.11%、0.66%となり、平均約19%上昇した。塩基サイトあたりのindelの数と長さを解析した結果(図1)、5' UTRへのindelの入り方は、(特に長さについて)チンパンジー挿入・ヒト欠失の傾向が見られ、チンパンジーの方が5' UTRが長くなっていた。CDSでは状況と逆で、ヒトの方に新規アミノ酸配列の獲得がより多く認められるという結果になった。

(2) ADRB3のアルギニンがトリプトファンに変化した型は他の哺乳類にはみられない。exon1をチンパンジー20個体のゲノムDNAについて調べたところ、この変異に関してヘテロ接合が3個体のみみだされた。したがって、アルギニンのトリプトファンへの置換はヒトではじめて生じたのではなくヒト・チンパンジー系統で生じたものと推定される結果を得た。

D. 考察

本研究における、チンパンジーの遺伝子の全長解析により、indelにみられる種間差が明らかになった。このindelの計算処理については、まだ考慮の必要がある。これまでの他の研究グループによる論文でも統一的な扱い方の記載が無い。本研究では、便宜的に保存された領域の長さに対するindelの出現率を求めた。一般に欠失が挿入の機構に比較して考えやすい。従って、両種を比較し、挿入型配列をもって祖先型として、配列上の変化率を計算することも考えられる。しかし、正確には、たとえばゴリラなどの配列をアウトグル