

2004年5月27日
独立行政法人 理化学研究所
ネイチャー・ジャパン

チンパンジー22番染色体の解読が終了

- ヒトとチンパンジーでは何が違う? -

独立行政法人理化学研究所（野依良治理事長）が中心となって組織した国際チンパンジーゲノム 22 番染色体解読コンソーシアム(The International Chimpanzee Chromosome 22 Consortium) は、ヒト 21 番染色体に相当するチンパンジー22 番染色体の解読を終了しました。

ヒトとチンパンジーの形態的及び行動的な大きな違いは、今日まで大規模かつ系統的には研究されておらず、わずかな遺伝的差異に基づくものだと考えられてきました。2003年7月1日、国際コンソーシアムは、ヒトに最も近縁であるチンパンジーの22番染色体の33.5Mb（3350万塩基）に渡る全配列を、99.998%の高精度で決定し、情報解読に成功しました。完成した配列データは国際コンソーシアムのウェブサイトから利用できますが、国立遺伝学研究所 DDBJ など公的データベース (<http://www.ddbj.nig.ac.jp>) から公開されています。

今回の解析から、ヒトとチンパンジーの間には塩基置換（1.44%）以外に転移因子などの挿入や欠失にも多数の違い（68,000カ所）が存在し、さらに比較された200余の遺伝子のうち80%以上で、生産されるタンパク質の機能に何らかの相違を生ずると思われる違いが見い出され、さらにレトロ転移がヒトとチンパンジーの進化に異なる影響を及ぼしてきたことを示唆する結果を得るなど、今まで考えられていた以上の違いが両者にあることが明確になりました。

ゲノムの解読は、医学における画期的な新アプローチの導入につながると同時に生物学の他のアプローチと組み合わせ、私たちヒトについての総合的理解を可能にすることが期待されています。本研究成果の詳細は、英国の科学雑誌『Nature』（5月27日号）に掲載されます。

1. 背景

平成12（2000）年3月に東京で開催された類人猿ゲノムと脳研究の総合的研究に関する国際ワークショップ“Genes and Minds Initiative (GEMINI)-Workshop on Ape Genomics”が、今回の世界に先がけたチンパンジー22番染色体を解読する国際コンソーシアム*1結成の発端となりました。このワークショップにおいて韓国、中国、台湾の研究者が、類人猿のゲノム計画が日本を中心としてアジア主導となることを歓迎し、ヒト21番染色体*2解読のメンバーであったドイツチームとともに、コンソーシアムメンバーとなる経緯となりました。また、理研GSCを中心とするグループは2002年1月にチンパンジー全ゲノムのBAC地図作成を完成し、ヒトとチンパンジーには1.23%の塩基の違いがあることを初めて明らかにしました

（Science誌発表）。その後コンソーシアムは22番染色体シーケンスを本格的に開始し今回の解読完了に至りました。

2. 研究の進め方

プロジェクトは以下に記した 9 機関によって進められました。配列決定は 8 機関が領域を分担し、2002 年 1 月から 2003 年 5 月にかけて全領域の解読を終了しました。シーケンス材料^{*3} (BAC、フォスミドライブラリ) の準備と比較地図の作成、連続クローンの作成と配列決定用クローンの供給及び進捗状況の管理は理研グループがプロジェクトを通じて担当しました。配列データは 99.998% 以上の高精度であり、22 番染色体長腕の 99% 以上が解読されました。期間中に何度かの全体集会 (上海、ベルリン、横浜) を開催してプロジェクトの進捗状況に関する調整を行ったほか、配列データ解析ではヒト 21 番染色体との比較、遺伝子同定、構造の特徴などの解析がコンソーシアム全チームの共同作業によって行われました。なお配列データの精度管理、データ解析、論文作成等プロジェクト全体を理研グループが主導しました。

3. 解析の手法

国際コンソーシアムは 3 個体のチンパンジー (日本 1 個体、アメリカ 2 個体) から調製したゲノムをシーケンスの材料としました。これら 3 個体の BAC ライブラリーから 22 番染色体に由来するクローンをスクリーニングし整列化を行いました。さらに理研グループによって作成された独自のフォスミドライブラリーもシーケンス未決定領域をカバーするため使用しました。シーケンスはクローンごとのショットガン法によって行い、クローンごとの配列をつなぎ合わせることで染色体全体の配列を編集しました。配列データはその精度が 99.99% 以上となるまで精査し最終的に高精度な完成配列を得ました。配列データは主に医学、進化、構造等の面からとくにヒトとチンパンジーの違いを知るためヒト 21 番染色体との比較を中心にコンピュータを用いて解析したほか、オランウータン、ゴリラといった類人猿ゲノム DNA との実験的な比較も行われています。

また、遺伝子については他の生物種 (マウスなど) との構造比較も行われたほか、ヒトとチンパンジー間での遺伝子発現量の実験的比較も行なわれています。今回の解析から、ヒトとチンパンジーの間には塩基置換 (1.44%) 以外に転移因子などの挿入や欠失にも多数の違い (68,000 カ所) が存在し、さらに比較された 200 余の遺伝子のうち 80% 以上で、生産されるタンパク質の機能に何らかの相違を生ずるとされる違いが見い出され、さらにレトロ転移^{*4}がヒトとチンパンジーの進化に異なる影響を及ぼしてきたことを示唆する結果を得るなど、今まで考えられていた以上の違いが両者にあることが明確になりました。

4. 今後の展開

今回の成果は、ヒト以外の霊長類で全配列の解読が終了した初めての染色体であり、今後はこれら成果をもとにヒトの進化や生物学的特質についての重要な情報がつぎつぎ得られるものと考えられます。ゲノム構造を基盤とする大規模研究は、医学における画期的な新アプローチの導入につながると同時に生物学の他のアプローチと組み合わせ、私たちヒト、さらには地球上の生命についての総合的理解を可能にすることと期待されています。

(問い合わせ先)

独立行政法人理化学研究所 横浜研究所
ゲノム科学総合研究センター

センター長 榎 佳之

Tel : 045-503-9151 / Fax : 045-503-9150

(報道担当)

独立行政法人理化学研究所 広報室

Tel : 048-467-9272 / Fax : 048-462-4715

Mail : koho@riken.jp

<補足説明>

※1 国際チンパンジーゲノム 22 番染色体解読コンソーシアム

2000年3月東京で行われた GEMINI ワークショップ (Genes and Minds Initiative Workshop) を機会に結成され、ヒト 21 番染色体に対応するチンパンジー 22 番染色体の解読を通して、ヒトゲノムの進化のメカニズムの解明を目指す。

※2 ヒト 21 番染色体

ヒトの最小染色体で、日本を中心とする国際チームによって 2000 年 5 月にその全配列が決定、解読された。ヒト染色体の中で最も深く理解された染色体の一つであり、遺伝子の高密度、低密度領域の存在、多様な転移因子の分布、ダウン症など様々な疾患に関する遺伝子の存在など、小さいながらもヒトゲノムの特色をそなえており、チンパンジーとの比較解析から得られる情報も多いと期待される。

※3 シークエンス材料

細胞・組織やゲノム DNA よりプラスミド、 λ ファージ、フォスミドまたは BAC など様々なベクターへクローニングを行い、目的にあったライブラリーに使われる材料。

※3 レトロ転移

「特定の転位因子の増殖」もしくは、「特定の反復配列の増幅」

主たるコンソーシアムメンバー

日本

榎 佳之 (理研ゲノム科学総合研究センター長)

渡辺日出海 (理研ゲノム科学総合研究センター・奈良先端科学技術大学院大学助教授 (現北海道大学教授))

藤山秋佐夫 (理研ゲノム科学総合研究センター・国立情報学研究所教授)

服部正平 (理研ゲノム科学総合研究センター・北里大学教授)

- 齊藤成也 (国立遺伝学研究所教授)
- ドイツ
- M.-L. Yaspo (マックスプランク研究所)
- Hans Lehrach (マックスプランク研究所)
- Svante Paabo (マックスプランク研究所)
- 中国
- S.-Y. Wang (中国上海ヒトゲノムセンター)
- Chen Zu (中国上海ヒトゲノムセンター)
- 韓国
- H.-S. Park (韓国基礎生物工学研究所)
- 台湾
- S.-F. Tsai (国立陽明大学)

チンパンジー22番染色体シーケンスと各チームの貢献度

| Center | Minimum tiling path clones | Additional clones | All clones | Total bases | Accession Numbers |
|---------|----------------------------|-------------------|------------|-------------|-----------------------------------|
| CHGC | 39 | 1 | 40 | 6,016,694 | BS000165-204 |
| KRIBB | 21 | 0 | 21 | 2,810,557 | BS000205-225 |
| Germany | 47 | 12 | 59 | 4,979,927 | AL954200-256, 258-259 |
| NIG | 5 | 2 | 7 | 1,050,886 | BS000158-164 |
| RIKEN | 143 | 14 | 157 | 19,540,921 | BS000001-157 |
| YMGC | 18 | 2 | 20 | 3,412,759 | BS000226-245 |
| Total | 273 | 31 | 304 | 42,666,429 | AL954200-256, 58-59; BS000001-245 |

Germany: MPI, IMB and GBF

CHGC: Chinese National Human Genome Center at Shanghai (China)

GBF: German Research Center for Biotechnology (Germany)

IMB: Institute for Molecular Biotechnology (Germany)

KRIBB: Korea Research Institute of Basic and Biotechnology (Korea)

MPI: Max-Planck Institute (Germany)

NIG: National Institute of Genetics (Japan)

RIKEN: RIKEN Genomic Sciences Center (Japan)

YMGC: National Yang-Ming University Genome Center (Taiwan)



協力したチンパンジーのゴン
(京大霊長類研究所)

チンパンジー22番染色体
及びヒト21番染色体シーケ
ンスの統計比較データ

| | ヒト21番染色体長腕 | | チンパンジー22番染色体長腕 | |
|---------------|------------|--------|----------------|--------|
| サイズ(塩基対) | 33,127,944 | | 32,799,845 | |
| 非対応サイト数*2 | 25,242 | | 101,709 | |
| 配列ギャップ数 | 14 | | 22 | |
| クローンギャップ数*3 | 3 | | 2 | |
| 推定クローンギャップサイズ | 73,108 | | 74,311 | |
| GC含量(%) | 40.94% | | 41.01% | |
| CG数 | 361,259 | | 358,460 | |
| CpG島数 | 950 | | 885 | |
| 塩基多様性 | 0.072% | | 0.14% | |
| 反復配列 | 塩基対 | 頻度 | 塩基対 | 頻度 |
| SINEs | 3,649,153 | 15,137 | 3,614,825 | 15,048 |
| Young Alus *4 | 21,557 | 75 | 2,606 | 10 |
| LINEs | 5,853,821 | 8,737 | 5,736,911 | 8673 |
| Young L1s *5 | 82,493 | 48 | 78,657 | 55 |
| LTRs | 3,621,501 | 7,282 | 3,550,807 | 7,180 |
| Transposons | 949,215 | 3,363 | 945,129 | 3,350 |
| RNAs*6 | 8,830 | 100 | 8,722 | 99 |
| Satellite | 19,327 | 21 | 14,773 | 18 |
| その他 | 30,452 | 38 | 34,776 | 43 |
| 計 | 14,132,299 | 34,678 | 13,905,943 | 34,411 |
| | 42.7% | | 42.4% | |

*1 Size of the contig data after the site where the first base of the PTR22q contig is aligned

*2 Regions extended into HSA21q clone gaps and subtelomeric unmatched regions

*3 Excluding pericentromeric and subtelomeric gaps

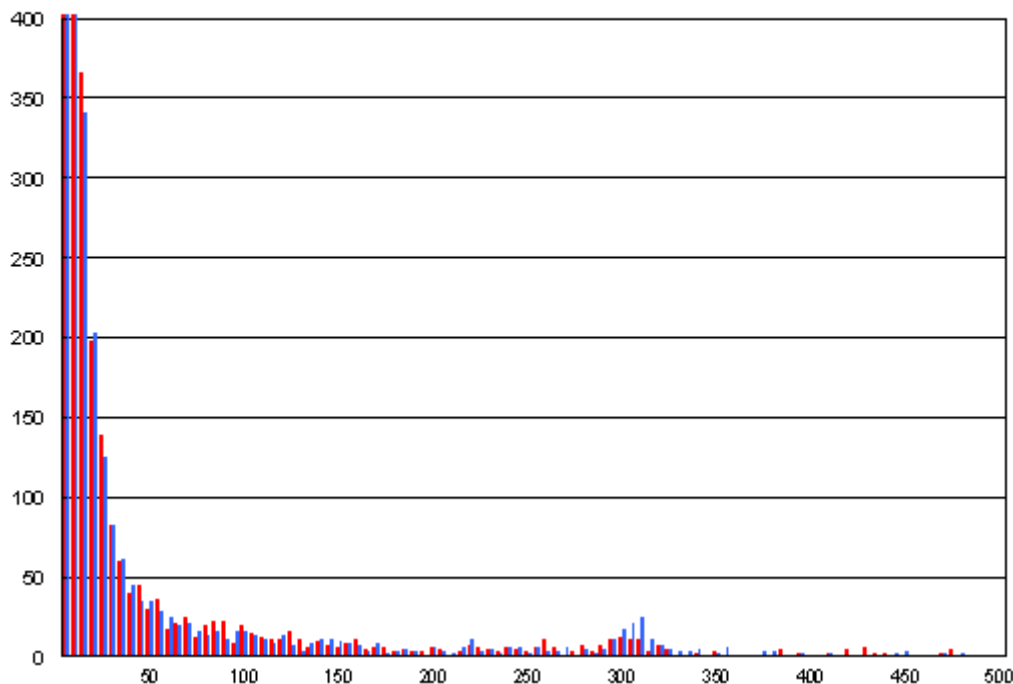
*4 AluYa5, AluYa8, AluYb8 and AluYb9

*5 L1HS and L1PA2

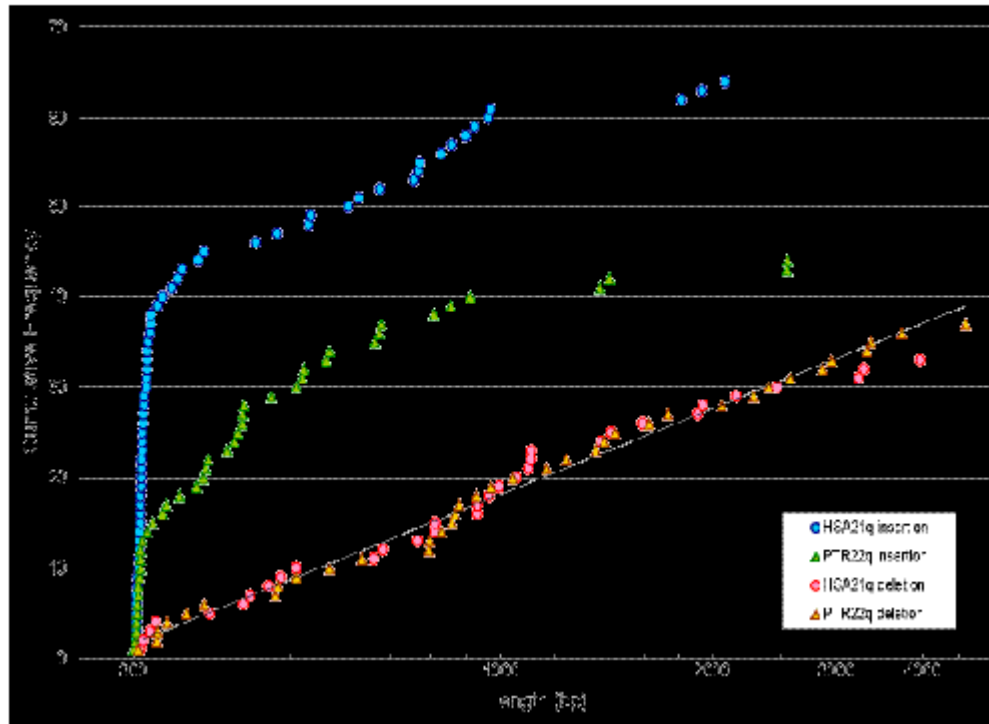
*6 snRNA, scRNA, 5S rRNA, tRNA, 7SL RNA and other small RNA genes

挿入・欠失のサイズと頻度

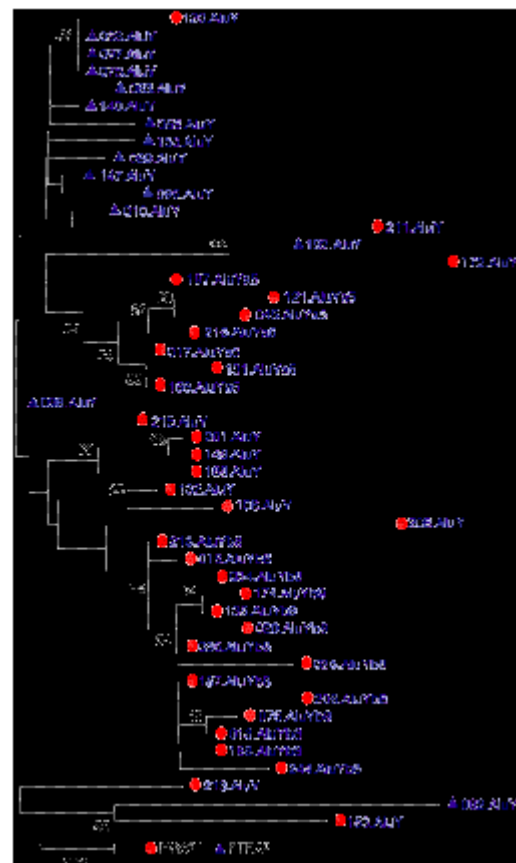
(Indelの総数: 約68,000ヶ所)



挿入 / 欠失頻度とサイズの関係



ヒト21番及びチンパンジー
22番染色体中にごく最近
挿入した転移因子Aluの進化的関係



Ka/Ks > 1: 進化速度の早い遺伝子

| | | |
|------|-----------|--|
| 3.37 | KRTAP23-1 | keratin associated protein 23-1 |
| 2.78 | C21orf87 | chromosome 21 open reading frame 87 |
| 1.98 | C21orf81 | chromosome 21 open reading frame 81 |
| 1.79 | C21orf128 | chromosome 21 open reading frame 128 |
| 1.76 | C21orf119 | chromosome 21 open reading frame 119 |
| 1.73 | RPS5L | ribosomal protein S5-like |
| 1.67 | KRTAP15-1 | keratin associated protein 15-1 |
| 1.57 | KRTAP21-1 | keratin associated protein 21-1 |
| 1.47 | ABCC13 | ATP-binding cassette, sub-family C (CFTR/MRP), member 13 |
| 1.37 | ANKRD21 | Expressed in prostate, ovary, testis, and placenta |
| 1.15 | TMPRSS2 | transmembrane protease, serine 2 |
| 1.01 | DSCR6 | Down syndrome critical region gene 6 |

進化速度の比較的早い遺伝子

| | | |
|------|-----------|---|
| 0.72 | TCP10L | t-complex 10 (mouse)-like |
| 0.56 | KRTAP13-3 | keratin associated protein 13-3 |
| 0.56 | KRTAP6-3 | keratin associated protein 6-3 |
| 0.54 | IGSF5 | immunoglobulin superfamily, member 5 |
| 0.35 | UMODL1 | uromodulin-like 1 |
| 0.34 | B3GALT5 | UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 5 |
| 0.32 | KCNE1 | potassium voltage-gated channel, Isk-related family, member 1 |
| 0.27 | KRTAP19-7 | keratin associated protein 19-7 |
| 0.19 | ATP5J | ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, |

ヒトとチンパンジーで
発現量の異なる遺伝子

| Gene ID | Transcript ID | Expression | | In age-specific changes | |
|----------|--------------------|-----------------|-----------------|-------------------------|---------------|
| | | brain | liver | brain | liver |
| C21orf18 | NM-017438 | higher in chimp | not expressed | chimp | |
| C21orf33 | Y07572 | not expressed | higher in human | | nonspecific |
| C21orf5 | AJ237839 | higher in human | higher in human | | chimp |
| C21orf97 | BC003651, AK024977 | higher in human | | nonspecific | |
| CRYAA | U05569, U66584 | higher in human | higher in chimp | | chimp & human |
| CRYLL1 | AK001293 | higher in human | not on chip | ? | |
| CXADR | Y07593 | higher in human | higher in human | | human |
| DSCR1 | U85267 | higher in human | higher in human | | human |
| ETS2 | J04102 | higher in chimp | higher in human | chimp & human | |
| IFNAR2 | X77722 | higher in chimp | higher in chimp | chimp | chimp & human |
| IFNGR2 | BC003624 | higher in chimp | | chimp & human | |
| ITSN1 | U61166 | higher in human | higher in chimp | | nonspecific |
| ISS | AK092334 | higher in human | higher in human | | nonspecific |
| EDXK | BC000123 | higher in human | higher in chimp | | chimp |
| PTTGIP | Z50022 | higher in human | higher in human | | chimp & human |
| TTC3 | D84296 | higher in human | higher in human | nonspecific | nonspecific |
| USP16 | AK025104 | higher in human | | nonspecific | |
| ZNF294 | AB018257 | higher in human | higher in human | chimp | chimp |

| |
|------------------------------|
| higher in chimp |
| higher in human |
| not expressed |
| no difference, but expressed |
| not on chip |

ヒトとチンパンジーの間で発現の違いや高いKa/Ksを示す疾患・感染症関連遺伝子の例

| Gene | Remark | Function |
|-----------------|-----------------------|--|
| ABCC13 | High Ka/Ks ratio | ABC transporter, ABCC13 is related with hematopoiesis |
| CRYZL1 | Expression difference | a quinone reductase similar to chloroquine-binding proteins identified in malaria-infected erythrocytes |
| CXADR | Expression difference | a receptor for coxsackievirus and adenovirus |
| DSCR1 (Adapt78) | Expression difference | an inhibitor of the serine/threonine phosphatase, calcineurin which dephosphorylates the tau protein associated with the assembly of paired helical filaments in Alzheimer's disease |
| ETS2 | Expression difference | a broad-range transcription factor which affects on the transcription of the APP gene associated with the pathogenesis of Alzheimer's disease |
| IFNAR2 | Expression difference | a receptor involved in interferon alpha-mediated antiviral activities for hepatitis B and C |
| IFNGR2 | Expression difference | a receptor involved in interferon gamma-mediated defense immunity against mycobacterial infections |
| ITSN1 | Expression difference | a member of a conserved family of proteins involved in clathrin-mediated endocytosis by which internalization of virus particles such as influenza virus into cells is mediated |
| TMPRSS2 | High Ka/Ks ratio | a transmembrane serine protease overexpressed in a majority of prostate cancer cells |
| TTC3 | Expression difference | a protein containing tetratricopeptide repeat (TPR) motif and mapped in DSCR, of which the mouse ortholog is expressed in predominantly in the central nervous system |