

The chimpanzee and us

チンパンジーとわれわれ

Wen-Hsiung Li & Matthew A. Saunders

チンパンジーゲノム概要配列の発表は、特別に重要なできごとだといえる。このゲノムは、ヒトの生物学的性質や進化を解明するための情報の宝庫となるからだ。

Nature Vol. 437(50-51)/1 September 2005

われわれヒトが系統的に最も近縁なチンパンジーとこれほど異なっているのは、いったいどんな遺伝情報の変化のおかげなのだろうか。研究意欲をかきたてるこの問題を解こうと、何十年も前から取り組みが行われてきたが、*Nature* 2005年9月1日号に発表されたチンパンジーゲノムの概要配列¹はその解明に向けた大きな一歩となる。ゲノム解析の対象となった種はナミチンパンジー (*Pan troglodytes*) で、その姉妹筋にあたる種はピグミーチンパンジー (別称ボノボ、*Pan paniscus*) だけである (図1)。

今回の概要配列から、ヒトゲノムとチンパンジーのDNA塩基配列は2~3%しかちがわないことがわかる。とはいえ、それぞれのゲノムは約30億塩基対からなるのだから、塩基対にして数千万個ものちがいに相当する。重要な相違が何かを判断する1つの手だては、われわれホモ・サピエンスに特異的な進化上の変化を見つけ出すことだ。もう1つは、両者のゲノムの塩基配列に残る「正の自然選択」の刻印を探すことである。概要配列の論文¹および同時掲載の論文²⁻⁴では、この2つの手法や他の比較解析法を取り上げている。

塩基配列をつなぎ合わせて全ゲノム配列を完成させるには、塩基配列を幾重にも読み取らなければならない。今回のチンパンジーゲノム概要配列では、読み取った塩基配列の全長は全ゲノム量のおよそ3.5倍で、この数字はヒトやマウス、ラットなど他のゲノムの最初の報告例よりも小さい。それでも今回の概要配列は、チンパンジーとヒトの両ゲノムの総体的なちがいをみるのに大いに役立つ。この新しいデータから、両者のゲノムは塩基置換からみるとわずか1.23%しかちがっていないことがわかる。この数値は、ゲノムから

ランダムに選んだそれぞれ500塩基対ほどのわずか53領域から得た従来の算定値⁵と一致する。

両者の塩基配列の分岐 (相違) の程度はゲノムの領域によって差があり、これはおそらく、細胞分裂に際して変異が起こる率や、選択による制約、染色体対どうしの塩基配列交換 (「組み換え」という) の率が、領域によって異なるためである。分岐の程度が最も大きいのはY染色体で、最も小さいのはX染色体である。これは予想どおりで、Y染色体が雄にしかなく、雄は雌より生殖細胞系列の変異率が高いのに対して、X染色体は雌雄両方に運ばれるためだろう。

自然選択はおもにタンパク質レベルで働くとして一般に考えられている。このため、タンパク質コード領域での塩基の置換は通常、2通りに分類される。つまり、「同義置換」(アミノ酸を変化させない置換) と「非同義置換」(アミノ酸を変化させる置換) である。あるコード領域が選択による強い制約を受けている場合、非同義的な塩基置換率 (K_A) は同義的な塩基置換率 (K_S) よりも有意に低くなるはずである。すなわち、 K_A/K_S 率は1未満となるはずである。一方、ある遺伝子が選択により非常に弱い制約を受けるか持続的な正の選択を受けた場合、 K_A/K_S 率は1に近くなるか、1を超える可能性がある。

ヒトとチンパンジーで13,454対の遺伝子を比較すると¹、 K_A/K_S の平均は0.23となり、もっと限られたデータセットからこれまでに比較算定されたヒトとチンパンジー (0.63)⁶ やヒトとヒヒ (0.34)⁷ の値よりもずっと低い。この比率は、マウスとラットの比較から算定された値 (0.13) の2倍である。これはおそらく、霊長類のように個体群サイズが比較的小

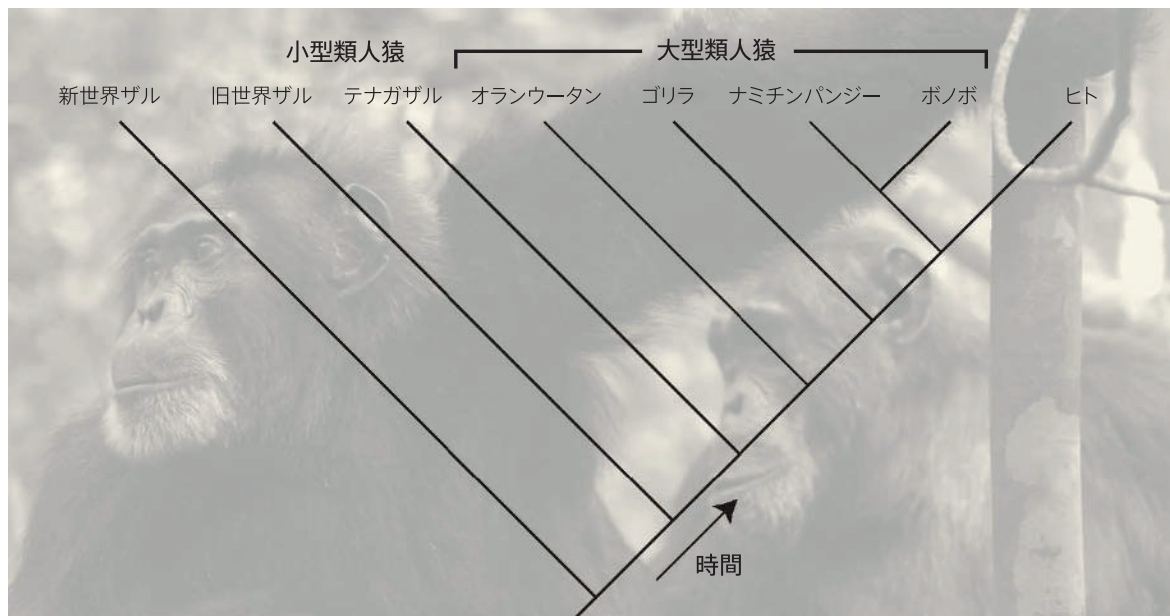


図1 高等な霊長類どうしの進化上の類縁関係。チンパンジー系統とヒト系統の分岐はおよそ600万年前に起こった。系統分岐の時間は線の長さに対応していない。

さい種では、浄化選択（有害な変異を排除する選択作用）があまり有効に働かないためだろう。重要なのは、この0.23という値が、ヒト間の多様性に関するデータから得られた K_A/K_S （およそ0.20~0.23）に近く、ヒト系統に生じた有利な変異の割合は従来の算定値^{7,8}よりも低いことが示唆される点だ。しかし、合計して585個の遺伝子（ランダムに起こるよりも多い）は、非コード塩基配列における塩基置換率（ K_I ）よりも高い K_A を示す。最も高い K_A/K_I の例としては、グリコフォリンCやグラニューリシン、プロタミン、セミノゲリンといった免疫や生殖に関係するタンパク質の遺伝子が挙げられる。

重複、挿入、欠失

塩基配列の分岐の程度を定量化する場合は単一塩基置換を考慮するのが一般的だが、DNA分節の挿入/欠失（略してインデル [indel] という）や最近起こった重複を考慮した場合、ヒトとチンパンジーのゲノムの相違はもっと大きくなる（それぞれ3%と2.7%）。挿入/欠失の3分の1をこえるものが反復配列のせいであり、およそ4分の1は転位因子（転位性遺伝因子）による。転位因子とは、別のゲノム領域に移動できるDNA配列のことで、主な種類としてAlu因子（長さ約300塩基対の短い転位性塩基配列）とL1因子（長い転位性塩基配列）の2群がある。

ヒトゲノムには7,000個ほどのAlu因子があるが、チンパンジーゲノムには2,300個ほどしかなく、チ

ンパンジーではAlu因子があまり活発に動かないことがうかがわれる。しかしL1因子は、従来の推定ではチンパンジーのほうが2~3倍活発だとされていた⁹が、今回、両者のゲノムで同じくらい活発なことがわかった。こうした相違が機能面にどんな重要性をもつのかは、まだ不明である。最近の分節重複（20メガ塩基以上の長さで配列同一性が94%をこえるもの）は、どちらのゲノムにも共通してみられる²。しかし、ヒトの重複分節のおよそ33%はヒトに特異的だが、チンパンジーの重複分節のうちチンパンジーに特異的なのは17%しかない。興味深いことに、ヒトに特異的な重複領域にある遺伝子のおよそ半数は、チンパンジーと比べて遺伝子発現にかなりのちがいがみられ、ほとんどの場合はアップレギュレーション（増産・機能亢進）を受けている。

ヒトの遺伝的多様性

チンパンジーゲノムの情報が得られたことで、ヒトの遺伝的多様性に関する既存の大量データを進化の研究にいかすことができる。これにより、ヒトの遺伝的多様性が大昔の祖先でどんな状態だったかを知ることが可能となり、ヒト集団内の遺伝子頻度データの助けを借りて、ヒトに最近（といっても25万年前よりも後のことだが）起こった正の選択の「痕跡」を見つけだせるかもしれない。選択的に中立であれば、新しい変異体が高頻度であられることはめったになく、種間の分岐の度合いは種内の遺伝的多様

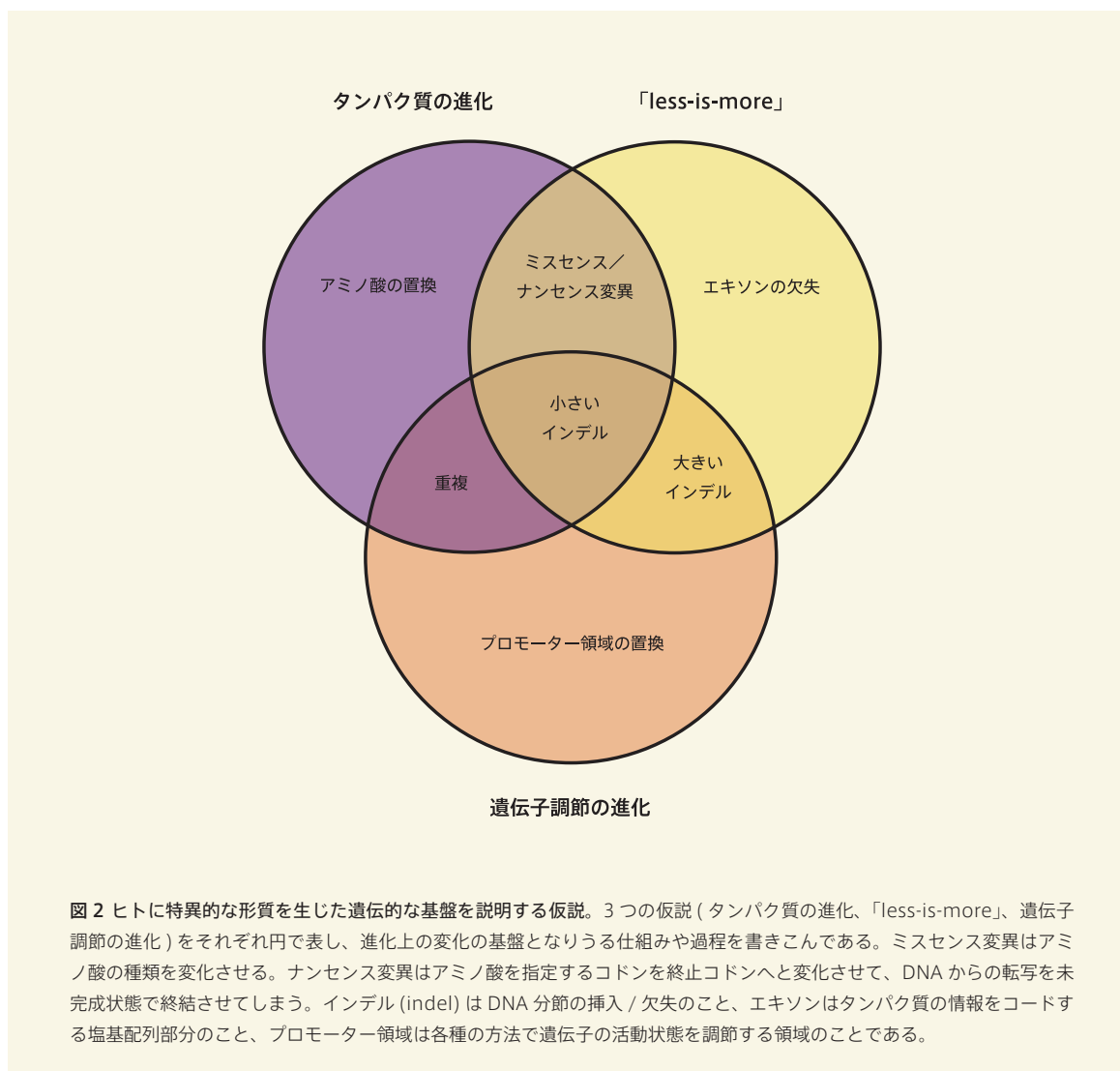
性の程度と相関しているはずである。現在の諸解析では、およそ 600 万年前に分かれたヒト属 (*Homo*) とチンパンジー属 (*Pan*) の 2 系統間の分岐の度合いから予想される値よりも変異が有意に少ないゲノム領域は、6 か所しか見つかっていない。つまり、これらの各領域はヒトで正の選択が最近働いたことを意味する。こうした手法の威力は、旧世界ザルやオランウータン（どちらのゲノム解析も現在進行中）などチンパンジーよりも遠縁にあたる霊長類のゲノム概要配列が完成すれば、一段と高まるだろう¹⁰（図 1、p11 News Feature 参照）。

ヒトをヒトたらしめるものとは？

どんな遺伝的変化のおかげで、われわれはヒトとして今ここにいるのかという問題は、上記の問題に比べるとはるかにやっかいだ。ヒトとチンパンジーの

ゲノムは非常によく似ているが、考慮すべき塩基の相違はおよそ 3500 万個、挿入 / 欠失が 500 万か所、そして多数の染色体再編成がある。これらの変化の大部分は生理的な影響をさほどおよぼさないだろうから、大きな脳容量や二足歩行、脳の発達（言語能力を含む）といった「ヒトらしさ」の特徴を生み出すゲノムの相違を突き止めることは、今なお気が遠くなるほどの難題である。ヒトとチンパンジーの系統が分かれてから間もないことを考えると、ヒトをチンパンジーや他の大型類人猿から区別する身体的な「表現型」の相違の一部は、影響の大きい少数の変異で生じたと思われる。

「ヒトらしさの形質」の進化を説明する説としては現在、有力なものが 3 つある。タンパク質の進化、「less-is-more」（遺伝子の減少が豊かさをもたらす）説¹¹、そして、遺伝子の活性を制御するゲノム領域の



変化¹²である(図2)。ヒトとチンパンジーの両ゲノムの予備解析から、これら3つの作用の相対的な寄与度を知るためのいくつかの手がかりが得られる。

では第1に、タンパク質の進化を考えてみよう。「ヒトらしさ」に貢献したアミノ酸変化は、急速に進化するタンパク質で見つかるのだろうか。 K_A/K_S が1をこえる遺伝子のほとんどは、想定されるヒトらしさの形質にかかわる生理過程に関与していない。実際、脳の機能やニューロンの活動に関係する遺伝子は K_A/K_S 値が平均に満たないのである。 K_A/K_S 値の高い遺伝子の大部分は、宿主-病原体相互作用や免疫、生殖に関するものだ。この傾向は、ラットやマウスその他の哺乳類でもみられる。このことから考えると、タンパク質の進化はヒト固有の形質の進化にさほど寄与していないのかもしれない。ただしこの可能性を捨てる前に、アミノ酸の置換を何度もこうむった遺伝子では、 K_A/K_S 分析で偏った値が出ることを念頭に置かねばならない。免疫や生殖に関与する遺伝子はとりわけ、こうした置換の影響を受ける。しかし、ほんの数回の変化の結果、これらの変化がきわめて有利なものであるために「選択による急速普及」(selective sweep; 選択圧により急速に集団内に広まること)をとげた遺伝子では、 K_A/K_S 値に重要な痕跡が残っていないだろう。たとえば、高度に保存されたFOXP2タンパク質(遺伝子転写因子の一種)に起こったわずか2つのアミノ酸変化が、ヒトの発話能力に関与した可能性もある¹³。結局のところ、ヒトとチンパンジーのタンパク質進化に挿入/欠失や遺伝子重複がどんな役割を果たしたかは、まだほとんど解明が進んでいない。

第2の「less-is-more」説では、ある種のヒトらしさの形質には、「類人猿の原型」の形質に機能欠損型の変異が起こったという特徴があると考えられる。例としてあげられるのは、体毛の薄さや、幼年期の形質の一部が成人になっても保存されていること、頭蓋骨の拡張などである。こうした機能欠損型の変異は、非同義置換や挿入/欠失、コード領域の消失、遺伝子まるごとの欠失によって起こりうるだろう。チンパンジーとの比較解析によって、コード領域に破壊的な挿入/欠失のあるヒト遺伝子が53個見つかり、こうした遺伝子をヒトらしさにかかわる興味深い表現型に結びつけることもできそうである¹⁴⁻¹⁶。挿入/欠失は、とくにこれらの変異がヒトらしさの進化について提案されている他の2つのメカニズムにも影響をおよぼしうることを考えると、ヒトとチンパンジーの表現型の相違に関与する大きな要因だと考えて差し支えないだろう(図2)。

第3に、ヒトとチンパンジーの表現型の相違は、主として遺伝子調節領域の変化から生じたとする長年の仮説がある。今回の解析¹では、この問題に詳しく取り組んではない。というのも、こうした領域を突き止めることは知っての通りまだむずかしいからだ。調節領域に関する現在の知識の大部分は、類縁関係の離れた生物種どうしの類似性を見つけたことから得たものである。この問題への取り組みは、ゲノム比較解析の枠組み内で、チンパンジーゲノム塩基配列との比較や機能を確認できるマイクロアレイ発現研究とともに、旧世界ザルなどの近縁な種どうしで保存されている調節領域を特定することによって¹⁷進めていけるだろう。遺伝子調節領域がらみの進化を説くこの仮説は現在、検証するのが最もむずかしい。だが、われわれが類人猿の生物学に対してヒトの生物学をどれくらい知っているかを鑑みれば、この仮説は最も期待がもてるかもしれない。

塩基配列の解読された脊椎動物ゲノムのリストに今回、チンパンジーゲノムの概要配列が加わったことは胸躍る出来事である。この概要配列は、ヒトの生物学や進化を解明するのにヒトゲノムそのものについて役立つ。しかし、このデータがあってもまだなお、ホモ・サピエンスと大型類人猿を隔てる主な特徴がいったいどんな遺伝的变化によって作りだされるのかについては、たくさんの疑問が残されたままである。この壮大なプロジェクトの次の段階では、個々の領域や遺伝子をさらに詳しく調べて、ゲノムレベルで今回明らかになった一般的なパターンを詳細に読み解くことが必要だろう。 ■

シカゴ大学(米)、Wen-Hsiung Li & Matthew A. Saunders

1. The Chimpanzee Sequencing and Analysis Consortium, *Nature* **437**, 69-87 (2005).
2. Cheng, Z. et al. *Nature* **437**, 88-93 (2005).
3. Linardopoulou, E. V. et al. *Nature* **437**, 94-100 (2005).
4. Hughes, J. F. et al. *Nature* **437**, 101-104 (2005).
5. Chen, F. C. & Li, W.-H. *Am. J. Hum. Genet.* **68**, 444-456 (2001).
6. Eyre-Walker, A. & Keightley, P. D. *Nature* **397**, 344-347 (1999).
7. Fay, J. C., Wyckoff, G. J. & Wu, C. I. *Genetics* **158**, 1227-1234 (2001).
8. Clark, A. G. et al. *Science* **302**, 1960-1963 (2003).
9. Mathews, L. M., Chi, S. Y., Greenberg, N., Ovchinnikov, I. & Swergold, G. D. *Am. J. Hum. Genet.* **72**, 739-748 (2003).
10. <http://www.genome.gov/10002154>.
11. Olson, M. V. & Varki, A. *Nature Rev. Genet.* **4**, 20-28 (2003).
12. King, M. C. & Wilson, A. C. *Science* **188**, 107-116 (1975).
13. Enard, W. et al. *Nature* **418**, 869-872 (2002).
14. Stedman, H. H. et al. *Nature* **428**, 415-418 (2004).
15. Hahn, Y. & Lee, B. *Bioinformatics* **21**, 1186-1194 (2005).
16. International Human Genome Sequence Consortium. *Nature* **431**, 931-945 (2004).
17. Boffelli, D. et al. *Science* **299**, 1391-1394 (2003).