

霊長類比較ゲノム解析とヒトへの進化機構の研究

◆藤山 秋佐夫¹⁾ ◆黒木 陽子²⁾

1) 国立情報学研究所生物系情報部門 2) 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＞

ゲノムの構造情報に基づき、種としてのヒトの特異性を明らかにすること、さらに、そこに至る進化機構を明らかにすることが本計画研究班の使命である。そこで、本研究ではまず、ヒトとヒト以外の霊長類間でゲノム構造の比較解析を行い、塩基配列、染色体構成の変化を明確にすることを目指した。これにより、ヒトを規定する上で意味を持つと予想される変化が見つけられることを期待してのことである。さらに、複数種の霊長類ゲノムを対象に相互比較を行うことにより、個別的検証が可能な数にまで対象領域を絞り込むことを、本研究班全体の計画とした。

本研究における藤山の役割分担は、まず研究遂行に必要な、細胞株、DNA試料、十分な被覆度を持つBACライブラリ、フォスミドライブラリ、染色体特異的ライブラリなどについて、外国からのリソースに頼ることなく自己開発した技術と材料を研究資材として広く利用可能な状態にすることである。

＜研究開始時の研究計画＞

ヒトゲノムの構造解析は着実に進行しており、それに伴って大量の配列データが蓄積されている。しかし、配列データのみから特にヒト特異的な構造上の特徴や機能を予測することは極めて困難であり、そのための有力な手段となるのが比較ゲノム解析である。マウスゲノムDNAの構造決定が進めば、両者の比較を行うことにより、哺乳類としての基本ゲノム構成が明らかになることが期待される。一方、ヒトに近縁な霊長類ゲノム構造との比較ゲノム解析を行うことにより、ヒトゲノムを特徴づける構造情報を明らかにすることも期待できる。そのためまず、ヒトゲノムの構造解析を実行し、そのなかでも特に解析の進んだ領域を対象として、それに相当する霊長類ゲノム（特にチンパンジーゲノム）との間で一次配列レベルでの比較構造解析を網羅的に行う。さらに他の霊長類ゲノムとの比較を行い、置換、挿入・欠失、転移などのゲノム構造上の変化と進化系統との関連を明らかにすることをめざす。

また、ゲノム研究に必要なライブラリと細胞株などの研究資材の整備を系統的に進める計画とした。

以上が研究開始時の状況をふまえた、当初の研究計画である。その後、動物ではヒトゲノム完全配列、チンパンジー、マウス、ラット、イヌ、トリ、トラフグ、ミドリフグ、ホヤなど急速に概要配列の発表が続き、当初に予想した比較ゲノム研究の重要性が、当初予想以上に大規模な実データを伴って証明されるに至っており、今後もこの傾向は一層の拡大が続けることが予想される。

＜研究期間の成果＞

ヒトゲノム構造解析

1999年12月の22番染色体、2000年5月の21番染色体に引き続き、我々は理化学研究所ゲノム科学総合研究センター、国際ヒトゲノムコンソーシアム（特に米国

Whitehead Institute）との協力のもとに、ヒト11番染色体、18番染色体を対象に構造決定を続けた。この成果は、2001年2月に国際ヒトゲノムコンソーシアムとしてヒトゲノム物理地図の決定とヒトゲノム概要配列の解析結果に関する初期解析として、論文報告されている（成果発表番号：1,2）。概要配列という精度上の限界はあるものの、ついに人類が我々自身の遺伝情報の全貌を見ることが可能になったことは、学術のみならず社会一般に与えた影響にも極めて大きなものがあつた。これらの成果が、20世紀から21世紀にかけて樹立された生命科学研究のマイルストーンとなることは間違いないだろう。

国際コンソーシアムが報告したヒトゲノムの概要配列は、物理地図情報に基づいて整理されたBACクローンごとに配列決定を行う、いわゆる階層的ショットガン法で決定されたものである（詳細は省略する）。データの大部分は多重度が5倍程度のショットガンデータに基づいたもので完成配列ではないが、真性クロマチン領域の90%程度（2.6～2.8 Gb）をカバーしており、塩基の読みとり精度については、91%の塩基が精度99.99%以上、96%が99.9%以上であった。既に完全配列が報告されていた染色体22番と21番を例外として、この配列は多数のギャップで分断されている。こうした性質のため、国際コンソーシアムでは概要配列と呼んでいた。また国際ヒトゲノムプロジェクトの成果として、塩基配列以外にも地図情報、マーカー情報、配列情報など、各参加機関のホームページや公共データベースで公開された情報に基づく解析結果も論文に含まれている。国際コンソーシアムの論文とほぼ同時に、セセラ社による全ゲノムショットガン法によるデータに基づいたヒトゲノムの概要配列も論文発表されている。全ゲノムショットガン法は精度を犠牲にしてスピードを重視した方法であり、現在では粗くゲノム全体を早く見るためには全ゲノムショットガン法が、より詳細な解析目的には高精度配列決定が行われる。日本は、大規模ゲノムを対象とした高精度配列決定が行える数少ない国の一つであり、ヒト全ゲノムの配列の完成配列化と、チンパンジーゲノム解析に大きな寄与を果たしている（後述）。また、全ゲノムショットガン法をヒトゲノムに適用したことの意味は別にもある。ヒトゲノム構造解析の全体的なスピードアップに大きく貢献したのは事実であるし、何よりも、強力な計算機パワーと優秀な情報科学者の組み合わせが、ゲノム研究には必須であることを明確に示したことが重要である。

ヒトゲノム情報は、別の意味でのインパクトも与えた。つまり、ここまでデータ量が膨大経済的にも重要な意味を持つようになると、利用形態が大きな問題となった。企業としてさまざまな付加価値を付けた有料サービスを行うのも一つの形態だが、WHOのヒトゲノムに関する宣言にもあるように、人類共通の資産として自由に利用されるべきである。国際ヒトゲノムコンソーシアムはデータの即時（24時間以内）公開を基本としており（いわゆるバミューダールール）、全データが国際DNAデータバンク（DDBJ/GenBank/EBI）に登録されている。登録デー

タについては自由に制限の付かない利用が可能である。

ヒトゲノム概要配列の遺伝子予測

概要配列論文の中で最もストライキングな内容は、ヒト遺伝子の総数についてである。従来の推定では、約10万程度とされていたのだが、我々は、染色体22番と21番の解析結果からヒトの遺伝子数を約4万と見積もった(図1)。その後、ESTデータの解析やフグの一種のゲノム解析結果に基づく遺伝子数の予測が相次いで発表され、一例を除いて3万〜5万の間に分布している。一方、ヒト遺伝子、遺伝病を集めたOMIMデータベースには、約1万個の遺伝子が登録されており、mRNAが収録されているRefSeqデータベース(NCBI)にもほぼ同数の配列が含まれている。

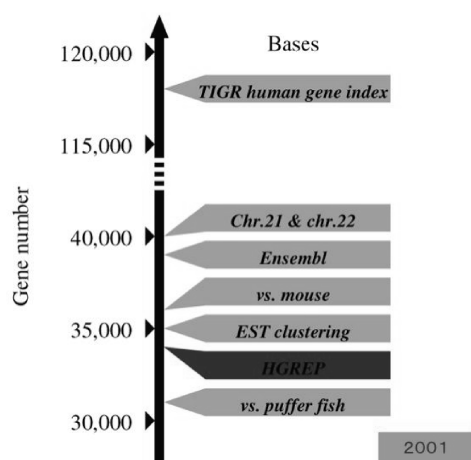


図1 ヒトゲノム概要配列に基づく遺伝子数予測

概要配列論文ではヒトの遺伝子数を30,000前後と推定していたが、この数値は完成配列の解析により、最終的に20,000-25,000に改められている。

ヒトゲノム完成配列の決定

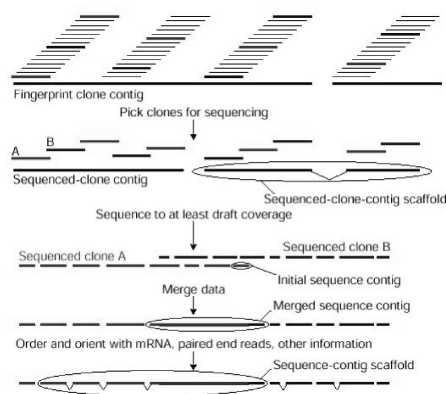


図2 完成配列の決定ストラテジー

概要配列の決定後もヒトゲノム国際コンソーシアムは配列の完成化作業を続行し、2003年4月にはヒトゲノム

完成配列の決定終了の宣言、さらに2004年に解析結果についての論文が報告されるに至っている(成果番号13)。基本的に行われた作業は、概要配列決定時に残っていたギャップ領域を徹底的に埋め、配列の高精度化を図ったことである(図2)。

最終結果は、図3に示すとおり、残存ギャップ数341、精度99.999%以上で決定された総塩基数が2,851,330,913 bpで真性クロマチン領域の99%をカバーするものであった(概要配列では、ギャップ数15万以上、精度99.9%、カバー率90%)。セントロメア領域、テロメア領域などのヘテロクロマチン領域のサイズを勘案すると、総ゲノムサイズが31億塩基対と推定された。また、予測遺伝指数についても、上述したように大幅な改訂が行われている。予測遺伝子数が減少した理由として、配列の高精度化と解析技術の進歩により、分断遺伝子の減少、偽遺伝子判別性能が向上したことが挙げられている。



図3 ヒトゲノム完成配列の解析結果

ゲノム研究リソースの系統的整備

ヒトゲノム：ヒトゲノム解読のために進めてきたリソース整備をさらに進めた。特に、我々が改良を続けてきた染色体ソーティング技術を駆使した、個別のヒト染色体、それを材料としたフォスミドライブラリ、BACライブラリはヒトゲノム解読に必要となった染色体の全てをカバーするものであり、多くの研究に利用されている(成果番号1,3,4,6-14)。

霊長類ゲノム：チンパンジー全ゲノムBACライブラリに加え、チンパンジー22番染色体、23番染色体、Y染色体特異的フォスミドライブラリとBACライブラリを構築した。また、ゴリラ全ゲノムフォスミドライブラリ、ニホンザルBACライブラリを構築し、オランウータンBACライブラリ構築の準備を進めている。また、ワシントン条約のために国内導入が困難になった霊長類細胞株の収集と樹立作業を進め(東京大学・石田貴文グループ、三和化研熊本霊長類パーク、京都大学霊長類研究所との共同研究、ナショナルバイオリソースプロジェクト、GAINグループの協力)、チンパンジー2亜種について、数株の細胞株を樹立した。我が国では、研究試料の用意は研究者個人の仕事であるという風潮が根強く、国の研究基盤としての研究リソースの重要性が認識され、上述したバイオリソースプロジェクト等への支援が始められるようになったのも始めたのはごく最近のことである。

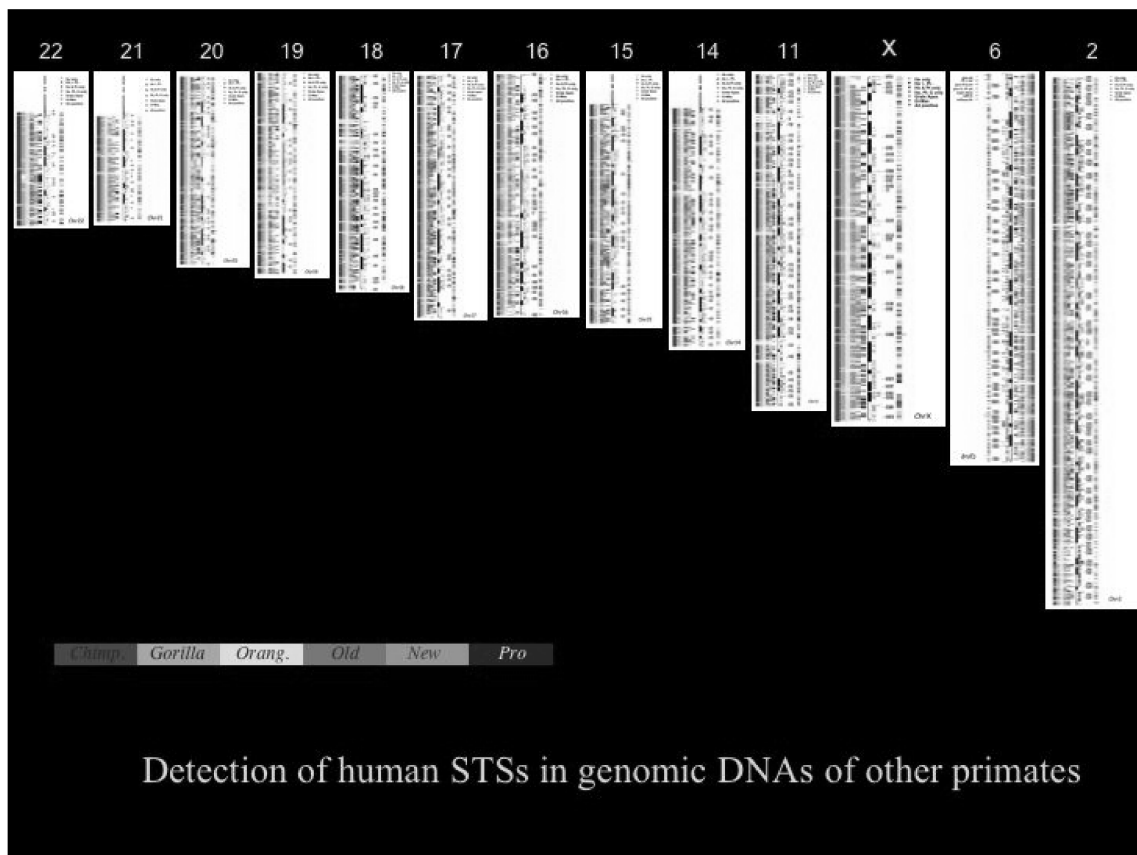


図4 ヒトゲノム配列をもとに、PCR法で比較した霊長類ゲノムの類似性。各染色体毎に、ほぼ等間隔に設計したプライマーを用い、チンパンジー（赤）、ゴリラ（オレンジ）、オランウータン（黄色）、カニクイザル（緑）、クモザル（青）、キツネザル（紫）細胞から調製したゲノムDNAを鋳型としてPCR反応を行った。反応生成物が確認できたSTS位置が着色してある。

ゲノム解析技術も重要な研究資源である。理化学研究所ゲノム科学研究センターと協力し、前述した染色体ソーティング技術、不死化細胞株樹立のほかに、FISH技術とファイバーFISH 技術の導入を行ったほか、以前から進めていたDNA コーミング技術に基づく単分子DNAファ

イバーの展開・検出とハイブリダイゼーション法の適用実験を行い、展開に関する条件設定を終了した。

霊長類との比較ゲノム研究

ヒトゲノムと、それ以外の霊長類ゲノム、特にチンパ

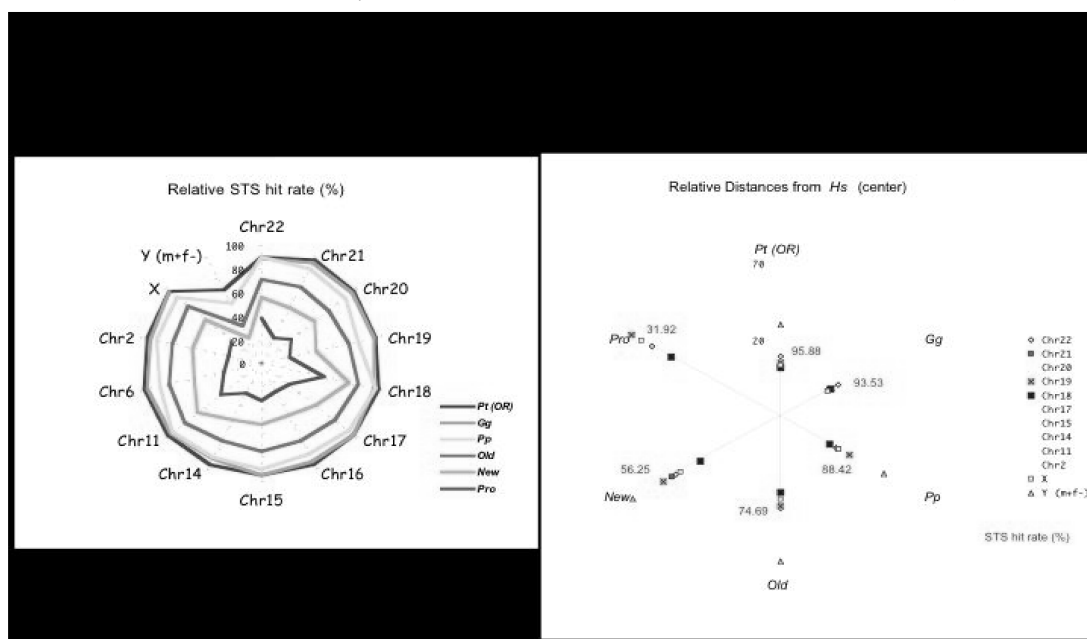


図5 ヒトゲノムに対する霊長類ゲノムの相対的類似性を示す。左図は、各染色体毎のPCR反応成功率。Y染色体マーカーの成功率が際だって低いこと、チンパンジーとゴリラのヒトに対する類似度が逆転している染色体の存在することがわかる。右図では、全STSについて、PCR成功率を種別に表した。例えばチンパンジーについては、挿入欠失まで含めた配列の類似度とPCRの成功率の値がほぼ近い値を示し、ゴリラとチンパンジーはオランウータンよりもヒトとの類似性が高いことを示唆する結果となっている。

Human STSs Contents in Various Primate Genomes

Chr	size(Mb)	STS	Hs+, Pt+	Hs only	Hs&Pt only	Hs, Pt, & Gg	Hs, Pt&Pp	Grate Apes	G+Mac	All
22	34.5	174	4	0	2	6	0	19	25	34
21	34	166	4	1	1	5	1	23	31	26
20	66	310	10	5	2	14	2	42	41	69
19	72	281	10	2	3	16	0	29	57	40
18	86	477	4	0	3	6	3	35	43	203
17	92	416	15	3	3	14	2	45	68	97
16	98	342	23	5	8	13	0	31	66	65
15	106	355	12	1	4	9	3	38	83	90
14	109	342	8	3	7	9	7	28	51	64
11	148	645	10	5	4	20	4	39	55	211
6	183	616	24	6	7	31	0	50	99	156
2	251	870	24	12	6	26	7	67	134	187
X	163	359	11	2	3	19	4	25	62	73
Y (Male+)*	60	32	7	3	3					
Total	1502.5	5385	166	43	56	188	38	471	815	1242
%	50.08		3.08	0.80	1.04	3.49	0.71	8.75	15.13	23.06

STS: effective STS

2001.12.6

* tested 65 STSs, 32 negative for female, 35 are positive for both male and female.

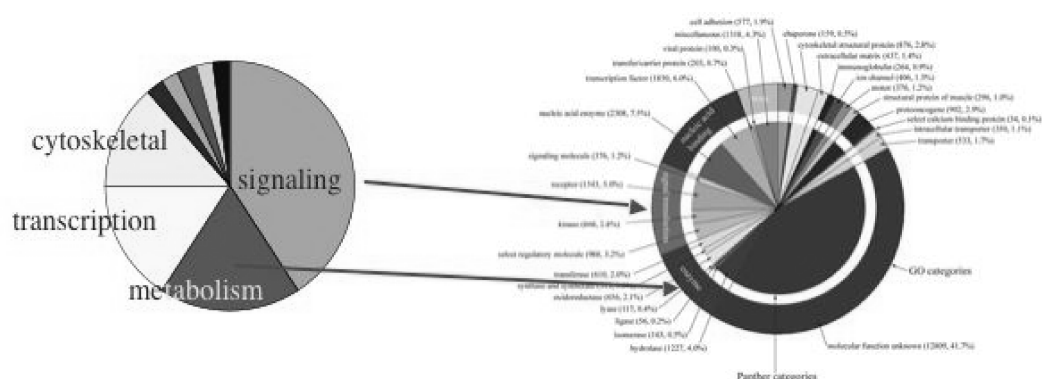


図6 ヒトSTSに基づく霊長類ゲノム比較

ンジーゲノムとの比較解析を始めるにあたり、まず大まかにゲノムの全体を対象とした比較を行う戦略を立てたが、共同研究先の理化学研究所がヒトゲノム解析に配列決定の努力を集中していたために、まず研究開始当時に利用可能であったヒト染色体配列（2,6,X,11,14-22）をもとに、比較的保存性が高いと考えられる遺伝子領域からほぼ等間隔（約200-300kb）にSTSマーカーを設定し、これらについて、ヒトゲノムDNAと、チンパンジー培養細胞株3種、ゴリラ、オランウータン、カニクイザル、クロオナガザル、ワオキツネザル培養細胞株から抽出したゲノムDNAを対象に、PCRベースで比較する方法を取った（藤山ら、未発表データ）。ヒト配列に基づくプライマーを用いたPCRの成功率を、相対的なゲノム配列の類似度と見なし、その結果を染色体上の位置に表示したのが図4である。かなり便宜的な方法だが、それでも、例えば染色体ごとにヒトゲノムとの類似性が異なることを示唆する結果や、霊長類ゲノム間で保存されているSTSや、いわゆる類人猿の中でのみ保存されているSTSが見つかるなど、ヒト以外の霊長類について全ゲノムにわたる配列情報を得ることが夢であった時代としては、まことに興味深い結果が得られていた（図5,6）。しかし、この手法から直接得られる情報はPCR反応が成功するか否かだけであり、最終的にはその領域に相当するクローンを単離して詳細な解析を行う必要がある。そのため、これと並行的に進めていたチンパンジー全ゲノムBACライブラリの構築（PTB1ライブラリ）を進め、最終的にこの計画は中断して、これとBAC-PAC Resourcesから導入したRP-43ライブラリからランダムに選んだクローンに対し、BAC-end配列の大規模配列決定と、ヒトゲノム配列へのマッピング、並びに配列比較を開始した。

ヒト-チンパンジー比較クローン地図の作成に基づく比較ゲノム研究

2001年2月に発表されたヒトゲノム概要配列についての歴史的論文は、「ヒトゲノムは、ヒトの発生、生理、医学、進化にかかわる膨大な情報の宝庫である。」と書き出される。また、ヒトゲノム計画の目標のひとつが、ヒトを構成する遺伝子などの遺伝情報の全体像と、その使われ方のすべてを明らかにすることであった。しかし、DNAの配列解読だけで遺伝情報の理解が充分行えるほどには、我々の技術も経験も進んでいない。したがって、いずれはそういう時代も来るのだろうが、当面は別の方法をとらざるを得ない。

そのため、生命情報学（bioinformatics）と比較ゲノム研究（comparative genomics）とが、ヒトゲノム計画の当初

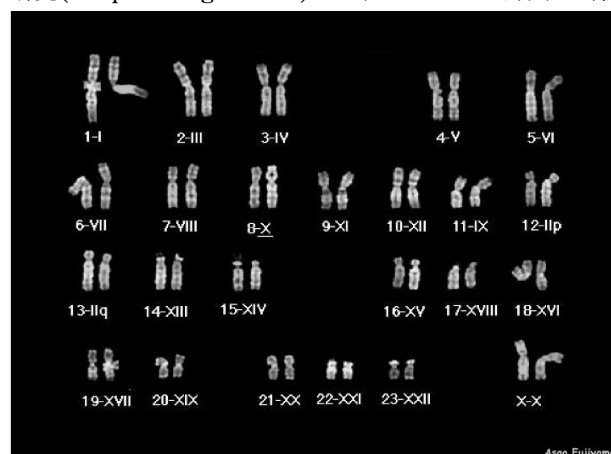


図7 ヒト（ローマ数字）とチンパンジー（アラビア数字）の染色体像

から着目されてきた。すなわち、ヒト個体間もしくはヒトとヒト以外のゲノムの構造を相互に比較、対照するという手法である。そのためにゲノム研究者が選んだ最初のは乳動物がマウスである。マウスは、ほ乳動物の代表として古くから研究の対象となっており、研究者の数も多いことから、ある意味では当然の選択といえるだろう。マウスのようにヒトと比較的進化的に離れた種類の場合には、比較するといっても2種類間での共通項を探すことになる。脊椎動物であり、胎盤を持つほ乳類であるというのが代表例である。もう一つ別の立場が、比較的近縁のものを比較し、両者間の違いを探すという方法である。正確な配列情報を元に精密な比較を行うことができれば、対象物の性質を際立たせることができ、共通点や相違点を明確に見られることが期待できる。

そのために我々が解析の対象に選んだのが、ヒトの最近縁種であるチンパンジーのゲノムである。進化上、ヒトとチンパンジーは数百万年前頃に分岐したとされており、歩行の形式、脳容積、体格、発声器官、生殖器官など、多くの表現型質の違いが知られている。これらの形質差のうちのどれがどの遺伝子に対応しており、また、ゲノムの構造（遺伝子構造）に違いがあるとするならば、それがどのような進化経路の結果生じたのか。ヒトをヒトたらしめた遺伝情報は存在するのか。また、存在するとしても、それ（ら）がどのような機能を持つのかといった問題が、そもそも問題として成立するのかということも含め多くの問題が研究の対象となる。

これまでにヒト以外の霊長類ゲノムを直接の解析対象とした研究は存在しなかったため、我々のチンパンジーゲノム研究は、まずゲノムの構造を比較するところから、そして、そのために必要な材料作りの部分から始めた。

Table 1. Summary of BES readings and mapping

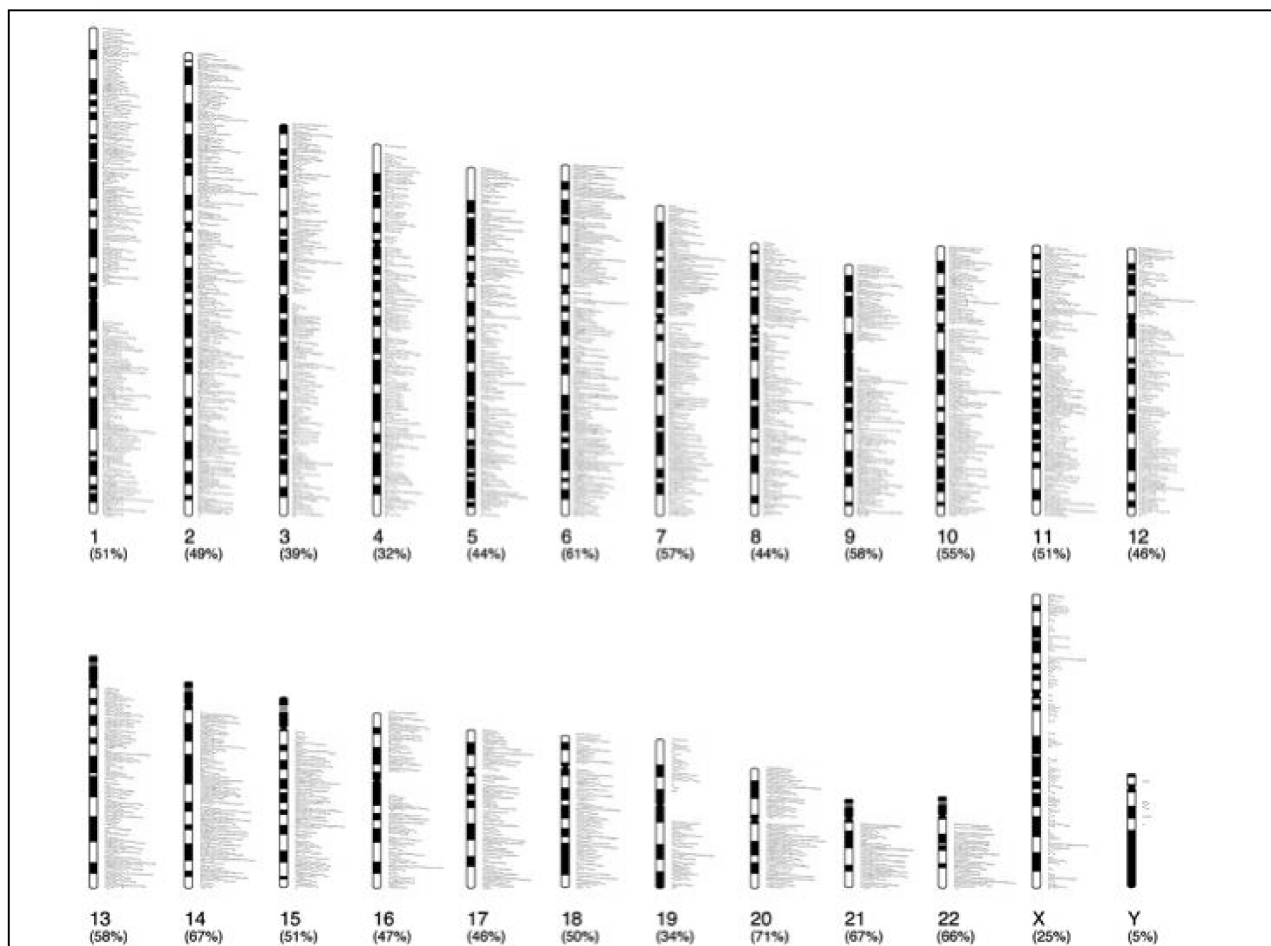
	Number of BAC Clones *	Number of BESs †
Sequenced	64,116	114,421
Mapped	52,881	77,461
paired ends ‡		49,160
single ends **		28,301
Unmapped		36,960
repeat s		1,168
human ††		20,376
non-human		515
no hit		14,901

* These clones were selected from the PTB1 or RPC1 -43 chimp anzee male BAC libraries

† Number of BAC -end sequences (BESs) used for the mapping .

我々以外にも、多くの研究グループがこの観点からの解析を既に始めているか、多大な興味を示している。また、ヒトとチンパンジーのゲノムの違いは塩基配列だけではなく、カリオタイプが違うこともよく知られている。すなわち、ヒトの2番染色体が、ヒト以外の霊長類では2本の中型染色体に分かれて存在する（図7）。ゲノムの内部には、この様に個々の塩基の違い以上に大きな規模での染色体構成の違いが存在する部位も多くあり、これらの違いが数多く組合わさった結果として、種の違いに発展する可能性も考えられる。これらは、種の分化、成立といった生物学研究の根幹に関わる問題でもあり、ヒトゲノム構造がより正確に分かると共に、チンパンジーゲノムの解析が進めば一層はっきりしてくることが当初の期待であった。

図8 ヒトーチンパンジー比較クロン地図（成果5より）



そこで、我々はまず単一個体のゲノムから作成したチンパンジーBACライブラリ（2種類）から約64,000クローンについて、インサート断片の両末端配列決定を行い、最終的に77,461本の配列をインフォマティクスを駆使してヒトゲノム配列上に整列させた（表1と図8）。こうして作成した整列クローン地図は、ヒトとチンパンジーゲノムの構造比較研究を進めるための第一歩となるものであり、ヒトの特定遺伝子領域に対応するチンパンジークローンが容易に単離できたり、ヒトとチンパンジーの間で大きく構造の変化している部分や配列の変化の大きい領域の探索に用いることが可能になる。この成果は、Science誌に論文発表した（成果5）、国内外に大きな反響を引き起こし「ゲノム特定領域」による2002年の一般講演の演題にも選択された。

BAC-end配列ごとにヒト配列に対する塩基配列の平均一致度を算出すると、98.77%という値が得られた（図9）。

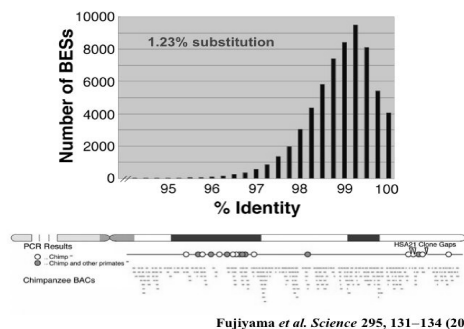


図9 チンパンジーゲノム断片のヒトに対する同源性分布

この値は、計算の元となった配列の総量が多いことと、全ゲノムにわたってランダムに配列決定が行われているため従来以上に正確な値であり、逆算するとゲノム中の1.23%、単純計算だと約3千7百万塩基について、ヒトとチンパンジーで塩基置換があるという結果である。2005年になって公表されたチンパンジー全ゲノムの概要配列解析からも全く同じ値が得られている。また、クローンPTB053J22については、ヒト12番染色体が分断された位置を含むことも明らかになった（図10）。

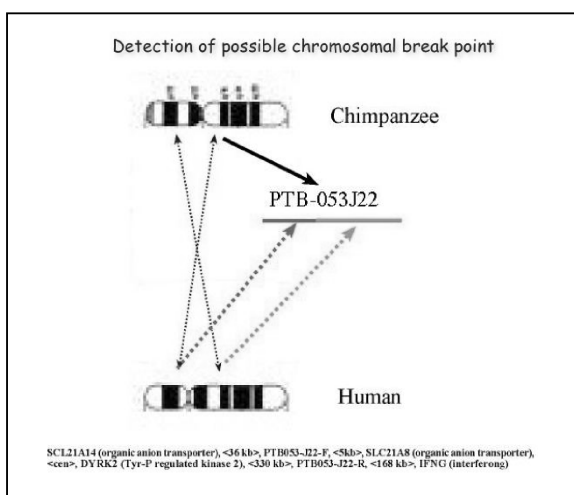


図10 ヒト染色体の分断位置をもつクローンPTB053J22

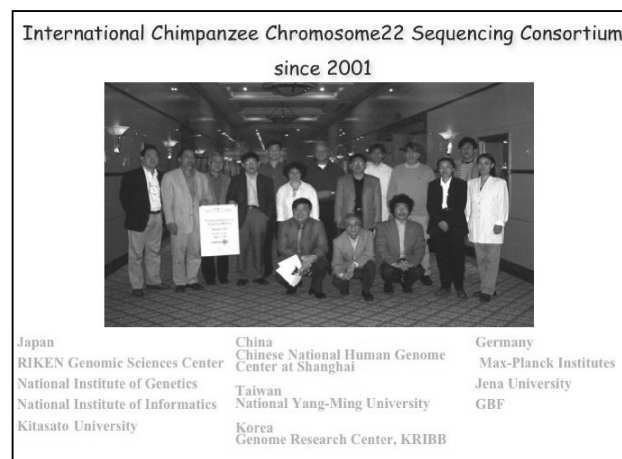
次に問題になるのは、この相違がゲノム全体に平均して分布するのか、それともある領域に局部的に集中して起きているのかという問題である（最終的には、両方のゲノムの全構造を比較すれば正確な解が出てくることになるが、チンパンジーゲノムの概要配列決定は大きく遅

れ、全容が発表になったのは2005年のことである）。そこでヒト21番染色体（チンパンジーでは22番染色体）についてBAC-end配列の分布を調べたところ、局部的に配列相同性の低い部分が集中していると思われる結果が得られた（図9）。また、末端配列をヒトゲノム上に染色体ごとに整列させ、その分布を調べたところ（図8）、Y染色体の配列に一致するものが他の染色体よりもかなり少ないことがわかった。Y染色体に関する同様な結果は、STSを基準にした比較解析でも同様な結果が得られており（図5）、Y染色体の進化速度が常染色体よりも大きいという従来の仮説をゲノムレベルでも示唆する結果であった。霊長類の種間で染色体構造を比較することからヒトの進化の道筋を染色体レベルでも明らかにすることは、今後の比較ゲノム研究課題の重要な課題の一つである。

チンパンジー22番染色体の高精度配列決定とゲノム比較

ここまでの段階で、ヒトゲノム上にチンパンジーゲノム由来のBACクロンの末端配列を配置し、各配列あたり数百bp程度の比較が行われた。しかし、両者の染色体間には転座や逆位のある部位も知られており、さらにこの手法では遺伝子に関する構造情報も断片的にしか得られないことから、より徹底してヒトとチンパンジーゲノムの比較解析を行う必要があった。そこで、ヒトの21番染色体とそれに相当するチンパンジー22番染色体の比較を、高精度配列に基づいて実施する計画が立てられた。そこで、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”に参加したメンバーを中心に、日本（理化学研究所、国立遺伝学研究所）、韓国（KRIBB）、中国（上海ヒトゲノムセンター）、台湾（国立陽明大学）、ドイツ（マックスプランク研究所）チームからなるチンパンジー22番染色体解読国際コンソーシアムが設立された（図11）。

図11 チンパンジー22番染色体解読国際コンソーシアム



この目的に利用できるチンパンジーの遺伝情報はほとんど存在しなかったため、日本のグループが中心となって配列を決定したヒト21番染色体の情報を徹底的に利用し、さらにチンパンジーについては個体の数を限定したライブラリを作成して実際の解析にあたった。我々が主に利用したのは、京都大学霊長類研究所で現在も生活しているオスのチンパンジー、ゴンのゲノムだが、他に2個体のチンパンジーゲノムから作成されたBACライブラリ（RP-43とCHORI-251）も利用した。また、BACライブラリに加え、22番染色体特異的のフォスミドライブラ

表2 ヒト21番染色体とチンパンジー 2 2 番染色体配列の統計値

	HSA21q	PTR22q
Size (bp)*1	33,127,944	32,799,845
Unaligned sites*2	25,242	101,709
# of sequencing gaps	14	22
# of clone gaps*3	3	2
Estimated total clone gap size	73,108	74,311
G+C%	40.94%	41.01%
CG dinucleotide	361,259	358,450
CpG islands	950	885
Nucleotide diversity	0.072%	0.14%
Repeats	bp	#
SINEs	3,649,153	15,137
Young Alus *4	21,557	75
LINEs	5,853,821	8,737
Young L1s *5	82,493	48
LTRs	3,621,501	7,282
Transposons	949,215	3,363
RNAs*6	8,830	100
Satellite	19,327	21
Others	30,452	38
Total	14,132,299	34,678
	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

リーもシーケンス未決定領域をカバーするため使用した。配列決定は2002年1月から2003年5月にかけて行われ、99.998%以上の高精度データで、22番染色体長腕の99%以上を解読した（クローニングギャップ2箇所）。シーケンス材料（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は、コンソーシアム全チームの共同作業によって行われている。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の運営は日本グループが担当した。

表2に、論文発表した配列解読の統計値をまとめた

（成果11）。このデータを元に計算すると、この2つの染色体間で観察された塩基置換の頻度は1.44%であったが、それ以上にさまざまな大きさの挿入・欠失部位の存在することが分かる（図12。LAGANプログラムを利用した解析結果）。配列をアラインメントし相互比較を行うことができた領域について挿入・欠失部位を数えると、その数は約68,000箇所確認でき、サイズ分布も1bpから40 kbを越えるものまで多様であった。図13に挿入・欠失部位のサイズごとの頻度分布を示したが、約300bp付近にピークの生じることが特徴的である（内容については後述する）。しかし、単純に配列を比較しただけでは、ヒト側のDNAに挿入が起きたのかチンパンジー側で欠失が起きたのかを区別することはできない（図14）。そこで、挿入欠失のうちで適当な大きさの部位を選び、挿入欠失領域の両端にPCRプライマーを設計して、ヒトとチンパンジ

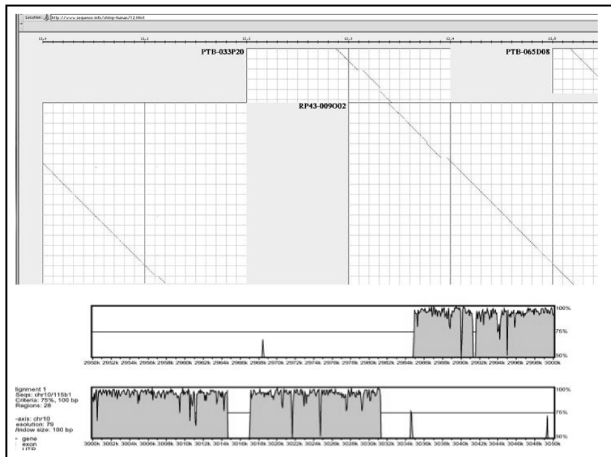


図 12 ヒト 21 番染色体とチンパンジー 22 番染色体の部分配列比較

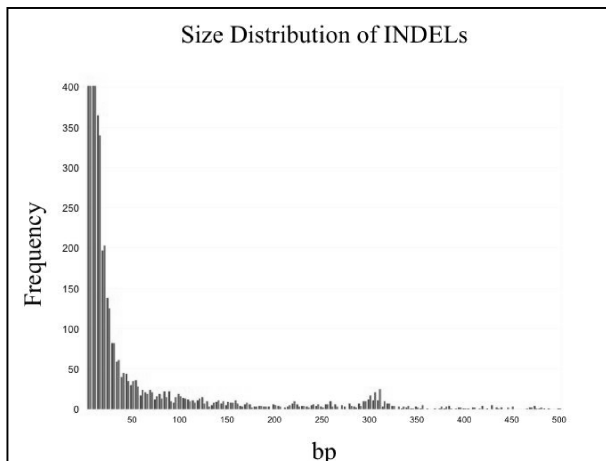


図 13 挿入欠失部位のサイズ別頻度分布

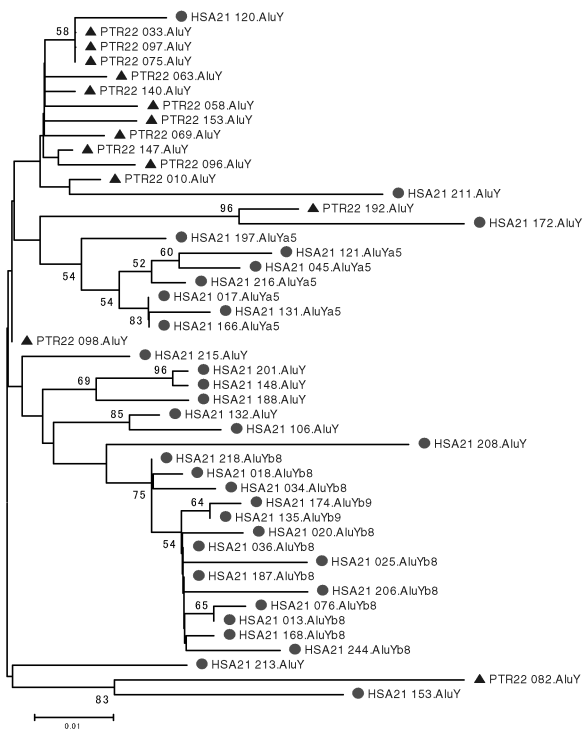


図 17 近隣結合法による、挿入 Alu 配列解析

簡単には区別が付かない

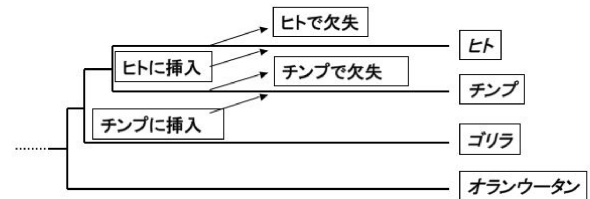


図14 種別に挿入欠失判定が必要

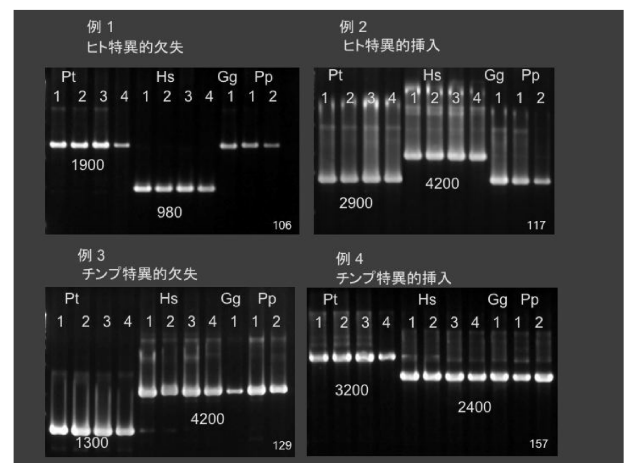


図 15 ヒト特異的挿入・欠失部位と、チンパンジー特異的挿入・欠失部位の検出

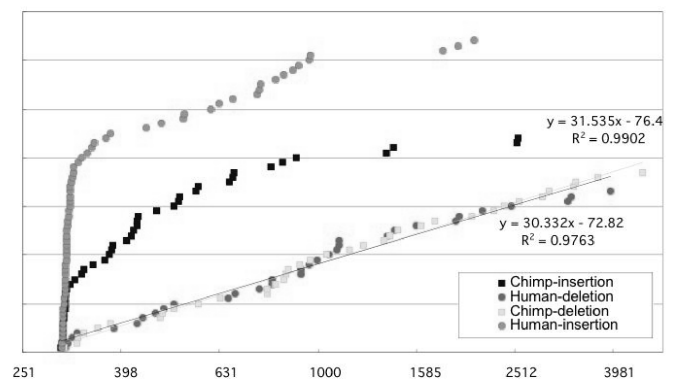


図 16 ヒト、チンパンジー別の挿入、欠失部位の度数分布。横軸は、挿入欠失のサイズ (bp)、縦軸は累積度数を示す。淡青：ヒト DNA に対する挿入頻度 濃青：チンパンジーDNA に対する挿入頻度 赤：ヒト DNA からの欠失 黄：チンパンジーDNA からの欠失

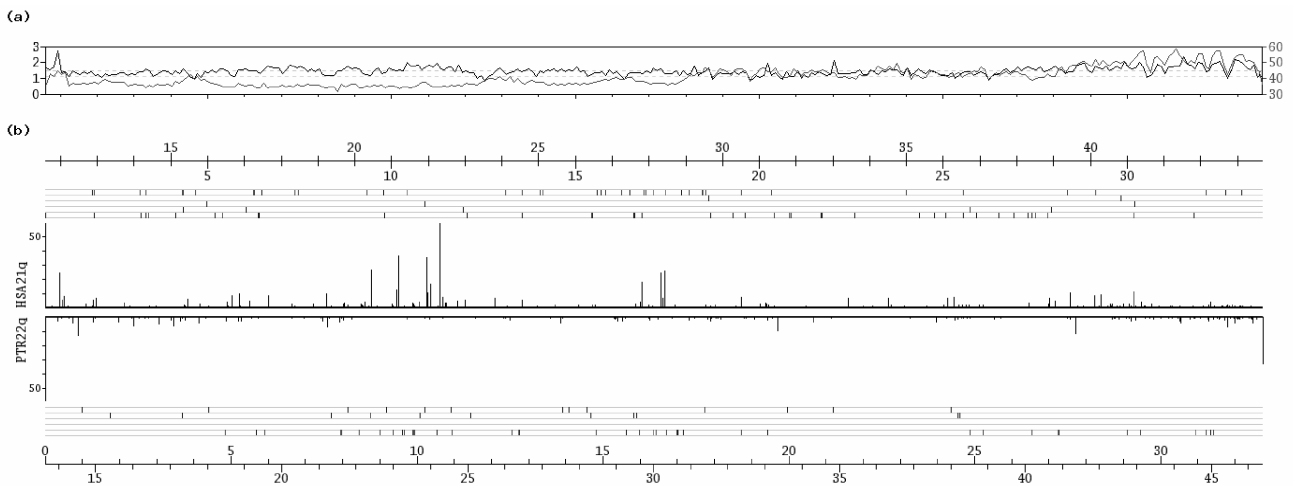


図18 ヒト21番染色体とチンパンジー22番染色体の配列比較。上段：黒線で塩基置換率、赤線でG+C%を示した。中段：ヒト21番染色体。からむ内は下からAlu, LINE, LTRの存在位置。下段：チンパンジー22番染色体。カラム内は上からAlu, LINE, LTRの存在位置。中央に、挿入欠失部位の位置とサイズを示した。

ー以外にオランウータンとゴリラのゲノムDNAについても検討を加えた（結果の一部を図15に示した）。この結果、部位によってはヒトとチンパンジーのそれぞれの系統で、挿入と欠失を明確に区別することが可能になった（図16）。図16に、ヒトとチンパンジー別に挿入、欠失の頻度をサイズに対して図示してある。この図から、ヒトではチンパンジーに比較して300bp程度の挿入が極めて高頻度に起きていること、また、欠失については生物種に関わりなく、サイズの対数に対してほぼ直線関係を示すことが分かった。こうしてはっきりと挿入部位であると判定できた配列のうち300bp付近の大きさのものはAlu反復配列であった。これらの相互関係を近隣結合法で比較したのが図17である。このデータの解釈については議論の余地がまだあるが、特にヒトにおいて、チンパンジーと分岐した後に、特定のAluサブファミリーが数度にわたって増殖、挿入している可能性が高いことを示している。遺伝子近傍に挿入が起きると、発現状態が変化したりする可能性も考えられるが、ゲノム全体についての精密な調査は今後の課題である。図18に、ヒト21番染色体とチンパンジー22番染色体の配列比較の結果を示した。

また、配列比較が可能な領域について、タンパク質をコードする遺伝子の構造比較を行ったが、約80%の遺伝子で何らかのアミノ酸置換が観察された。この差が機能上どのような意味を持つのかについては、今後の検討課題である。また、ドイツのグループにより、マイクロアレイを用いた遺伝子発現量についての比較も行われたが、結果の解釈については、やはり今後の検討課題である。

〈国内外での成果の位置づけ〉

以上の成果は、世界に先駆けてヒト以外の霊長類で全ゲノムを対象とした大規模解析を実施した初めての例であり、世界的にも大きなインパクトを与えたものと確信している。特に米国では、霊長類ゲノムとしてチンパンジーではなく、もっぱら医学実験に利用できAIDSも発症するという観点からマカクが選ばれていたのだが、米国を含まない国際コンソーシアムに解析を先んじられたことにより、結果的にチンパンジー全ゲノムの概要配列決定に踏み切ったともいえる。その結果、現時点でチンパ

ンジーはゲノム全体の概要配列と染色体特異的な高精度配列が利用可能な生物種となり、このことは特にヒトへの進化を考える上で極めて重要な意味を持つ。

また、学術面での影響のみならず一般社会に対しても、進化上の最近隣種のゲノムが決められ、特に塩基置換が1.23%程度の差しかなかった事実が、国内外のメディアで広く取り上げられた。ゲノム研究の科学としての面白さを市民に伝えることができたのであれば、我々としては望外の喜びである。

〈達成できなかったこと、予想外の困難、その理由〉

霊長類ゲノムの比較研究は、チンパンジーゲノム解析が予想以上に進んだ。この点での研究遂行上の最大の困難は、やはり理化学研究所に於いてヒトゲノム解読で培ってきたリソース、配列決定、インフォマティクスが融合したインフラが崩壊しつつあることである。したがって、国際協力、特にアジア地域を中心とした国際協力が今後も一層重要になることは間違いない。

技術開発については、DNAコーミングを担当していた学振特別研究員が任期途中で就職したため、特に最終的なハイブリダイゼーションに関する部分が条件検討を行ったのみで中断せざるを得なかった。

〈今後の課題〉

ヒトゲノム完全配列の報告以来、チンパンジー、マウス、ラット、イヌ、トリ、トラフグ、ミドリフグ、ホヤ（発表順不同）など急速にゲノム概要配列の発表が続いている。本研究の開始当初に予想した比較ゲノム研究の重要性が、当初の予想以上に大規模な実データを伴って証明されるに至っており、今後もこの傾向は一層の拡大を続けることが予想される。しかし、ヒトゲノム配列の完成宣言に伴い、それまで我が国を代表してヒトゲノム配列決定国際共同体に参加し、米英に次ぐ重要な地位を占めていた理化学研究所ゲノム総合研究センターヒトゲノム構造グループが理化学研究所内で予算と人員の両方で大幅な再編縮小を余儀なくされ、国内外における相対的な地位が低下するに至ったことは残念の極みである。米国、英国では、それぞれのゲノムセンターが多様性と比較ゲノムをキーワードに従来以上の活動を継続し、さらに韓国などヒトゲノム構造決定に参加していなかった国においてもゲノム解読の重要性が重視され、相応の施策がとられている。リソース、配列決定、情報処理のゲノム解読に重要な機能をバランスよく持っていた組織がこのような事態に至ったことは、我が国の基礎科学推進にとっては大きな痛手である。科学政策が目先の成果に振

り回されることのないよう、長期の視点にたった継続的戦略の立案が必要である。国際協調・国際共同研究は、同時にパワーゲームの場でもある。基盤が不安定では科学立国もおぼつかないだろう。

霊長類ゲノム研究については、我々の成果に加え、2005年のチンパンジー全ゲノム概要配列、Y染色体の構造解析論文と、相次いで成果が報告された結果、極めて多様な研究をさまざまなスケールと視点から行える状況が一応整ったといえる。したがって、今後の研究は徹底した情報解析にもとづいた、個別的な機能、進化研究と、適正なターゲットを狙った、比較的規模の大きい比較ゲノム研究の両面から研究を進める必要がある。

謝辞

本研究の共同研究者、特に試料調製、配列決定、配列情報処理等に、文字通りに骨身を惜しまぬ努力を払ってくれたスタッフに心から感謝する。

〈研究期間の全成果公表リスト〉

1) 論文

1. 111220124
The International Human Genome Consortium including A. Fujiyama and 55 authors: A Physical map of the human genome *Nature*. 409, 934-941 . (2001)
2. 111220127
The International Human Genome Consortium including A. Fujiyama and 72 authors from international sequencing centers: The Human Genome: Initial Sequencing and Analysis. *Nature*. 409, 860-921 . (2001)
3. 111220122
Thomas Br殫s, Gabor Gyapay, Jean-Louis Petit, Fran脛is Artiguenave, Virginie Vico et al. including Asao Fujiyama,: A clone map of human chromosome 14. (2001) *Nature*. 409, 947-948 (2001)
4. 111220130
Daniel Graf, Paula M. Timmons, Megan Hitchins, Vasso Episkopou et al. including Asao Fujiyama: Evolutionary conservation, developmental expression and genomic mapping of mammalian Twisted gastrulation. *Mammalian Genomics*, 12, 554-560 (2001)
5. 203032300
Asao Fujiyama, Hidemi Watanabe, Atsushi Toyoda, Todd D. Taylor, Takehiko Itoh et al: Construction and Analysis of a Human-Chimpanzee Comparative Clone Map. (2002) *Science*. 295, 131-134
6. 203032306
Yoshihisa Watanabe, Asao Fujiyama, Yuta Ichiba, Masahira Hattori, Tetsushi Yada, Yoshiyuki Sakaki and Toshimichi Ikemura: Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: disease-related genes in timing-switch regions. (2002) *Human Molecular Genetics*, 11, 13-21
7. 304281146
Masakatsu Watanabe, Naoki Kobayashi, Asao Fujiyama, Norihito Okada: Construction of a BAC library for *Haplochromis chilotes*, a cichlid fish from Lake Victoria. (2003) *Genes Genet. Syst.* 78, 103-105
8. 403311141
Kunihiko Takamatsu, Kohei Maekawa, Takushi Togashi, Dong-Kug Choi, Yutaka Suzuki, Todd Taylor, Atsushi Toyoda, Sumio Sugano, Asao Fujiyama, Masahira

- Hattori, Yoshiyuki Sakaki, Tadayuki Takeda: Identification of Two novel primate-specific genes in DSCR. (2002) *DNA Research* 9, 89-97403311153
- Construction of a BAC library for *Haplochromis chilotes*, a cichlid fish from Lake Victoria
9. 403311156
Dukka Bahadur K.C., Tatsuya Akutsu, Etsuji Tomita, Tomolazu Seki, Asao Fujiyama: Point Matching Under Non-Uniform Distortions and Protein Side Chain Packing Based on an Efficient Maximum Clique Algorithm. (2002) *Genome Informatics* 13, 143-152.
 10. 403311200
Choong-Gon Kim, Asao Fujiyama, and Naruya Saitou: Construction of a Gorilla Fosmid Library and its PCR Screening System. (2003) *Genomics* 82, 571-574.
 11. 407081936
The International Chimpanzee Chromosome 22 Consortium including A. Fujiyama: DNA sequence and comparative analysis of chimpanzee chromosome 22. (2004) *Nature* 429, 382-388.
 12. 502212103
Takashi Suzuki, Tatsuya Ota, Asao Fujiyama, and Masanori Kasahara: Construction of a bacterial artificial chromosome library from the inshore hagfish, *Eptatretus burgeri*: A resource for the analysis of the agnathan genome. *Genes & Genetic Systems*, (2004) 79, 251-253
 13. 502212126
International Human Genome Sequencing Consortium including Asao Fujiyama: Finishing the euchromatic sequence of the human genome. (2004) *Nature* 431, 931 - 945
 14. 503080806
Takashi Suzuki, Tadasu Shin-I, Asao Fujiyama, Yuji Kohara, and Masanori Kasahara: Hagfish Leukocytes Express a Paired Receptor Family with a Variable Domain Resembling Those of Antigen Receptors. *J Immunol*. 174, 2885-2891 (2005).
 - 2) データベース/ソフトウェア
なし
 - 3) 特許など
なし
 - 4) その他顕著なもの、
 15. 303271820
藤山秋佐夫：ヒトとチンパンジー ―ゲノムにかかれた進化の歴史―、*科学* 72, 389 (2002)
 16. 303271823
藤山秋佐夫：ゲノム地図におけるヒトとチンパンジーの違い、*実験医学*20、1011-1014 (2002) .
 17. 303271826
藤山秋佐夫：霊長類の比較ゲノム科学、*人工知能学会誌* 18, 20-23 (2003).
 18. 303271850
黒木陽子、藤山秋佐夫：ポストゲノムシーケンス時代の霊長類研究 *生体の科学* 53、566-571 (2002)
 19. 303271853
藤山秋佐夫：ゲノム時代の比較ゲノム学、*科学* 73, 407-411 (2003) .
 20. 305161750
渡邊日出海、藤山秋佐夫、服部正平、榊佳之：ヒト・チンパンジー比較ゲノム地図、*蛋白質核酸酵素*47、

808-813 (2002).

21. 403311207

藤山秋佐夫：ヒトゲノム計画の意義、成立と経過、学術月報 56, 781-785 (2003).

22. 403311209

黒木陽子、藤山秋佐夫：ゲノム解析技術とヒトとチンパンジーの比較ゲノム研究、システム／制御／情報 48, 80-85 (2004) .

23. 403311211

Shinichi Morishita and Asao Fujiyama: Body Expression Map of the Human Genome (2004) in Ency. of Molecular Cell Biology and Molecular Medicine (ed. Meyers, R.A.) Wiley-VCH, Germany