特集 霊長類学における行動・生態研究手法の現在

総説

霊長類分子生態学における次世代シーケンシング

早川卓志^{1), 2)}

1) 京都大学霊長類研究所ワイルドライフサイエンス（名古屋鉄道）寄附研究部門

2) 公益財団法人日本モンキーセンター

1. DNA を用いた霊長類分子生態学

直接観察だけではわからない、目に見えないものごとが、野生霊長類の行動や生態の背景に大きく影響していることが多くある。遺伝という現象もそのひとつだ。遺伝とは、親から子へ、環境要因に左右されることなく形質が受け継がれることをいう。形質とはあらゆる段階のことを指し、顔かたちがよく似るという形態レベルのこと、アルコールに強いとか乳糖が分解できるといった生理レベルのこと、性格や感覚が類似するという知覚・認知レベルのことなど、さまざま。

遺伝というマクロな現象を制御する分子の正体が DNA (deoxyribonucleic acid: デオキシリボ核酸) である。アデニン (A), チミン (T), グアニン (G), シトシン (C) という 4 種類の塩基を持つヌクレオチド分子が鎖状につながってできた高分子で、二重螺旋構造を形成し、細胞のひとつひとつに収納されている。このつながりはとてつもなく長く、ヒト (*Homo sapiens*) などの哺乳類ではおよそ 30 億 (3 ギガ) 個の異なる塩基を持つヌクレオチドから構成される (International Human Genome Sequencing Consortium, 2001; Kapusta *et al.*, 2017)。おのおの塩基配列は遺伝子として細胞の中で役割を果たし、さまざまな構造や機能を持つ数万種類のタンパク質をコードする。4 種類のヌクレオチドが 30 億個連なってできた哺乳類の塩基配列が持つ情報量は、4 の 30 億乗、すなわちおよそ 10 の 18 億乗という壮大な数である。こうした膨大な情報が DNA に刻まれているおかげで、生物は複雑な生命活動を制御できるのだ。

個々の生物が持つ塩基配列情報のすべてを、遺伝子 (gene) の全体 (-ome) という意味で、ゲノム (genome) と呼ぶ。ゲノムの塩基配列情報は、ほぼ正確に細胞分裂時に複製され、生殖細胞を介して、親から子へ受け継がれる。性を持つ生物では母と父の塩基情報が原則的に半

分ずつ継承される。これが親と子の間に遺伝という現象が見られる分子基盤だ。一方で、遺伝形質の個体差や種差が生じるためには、ゲノムの塩基情報に個体差や種差がなければならない。細胞分裂時の塩基配列情報の複製はほぼ正確であるが、一定の割合で変異が生じる。ヒトやチンパンジー (*Pan troglodytes*) であれば 1 世代あたり、およそ 1 億分の 1 の確率で塩基変異が生じる (Roach *et al.*, 2010; Tatsumoto *et al.*, 2017)。ゲノムは 30 億塩基あるので、1 世代あたり数十個の塩基変異が生じる計算になる。1 世代で見れば微小な個数であるが、1 万年も経てば数百は世代交代するので、その間に生まれる変異の蓄積が個体差や種差にのちととなることは容易に想像できる。例えば、およそ 10 万年前にアフリカから拡散したヒトのゲノムには約 0.1% の個体差、数百万年前に種分化したヒトとチンパンジー (*Pan troglodytes*) のゲノムの間には約 1% の種差が生じている (International Human Genome Sequencing Consortium, 2001, 2004; Chimpanzee Sequencing and Analysis Consortium, 2005)。

本稿では、野生霊長類の行動や生態を理解する先進技術としても活用されている DNA の研究についての歴史と現状について総説する。研究分野としては分子生態学 (molecular ecology) と呼ばれる領域に入る。分子生態学研究を可能としている DNA の 3 つの性質について、指摘しておきたい。

- ① DNA は細胞で機能するという制約があるため、原則としてどんな細胞にも含まれている。
- ② DNA の塩基配列はほぼ正確に細胞間で複製される。ただし微小な変異も生じる。
- ③ DNA は生物の形質を制御するタンパク質を遺伝子の形でコードする。

以上の DNA 特有の性質が、分子生態学研究でもたくみに利用されている。①の性質のおかげで、毛や糞など、どんな由来の細胞であっても、対象個体の DNA を調査

できることが保証されている。これは野生霊長類の調査の基本である非侵襲サンプリングで生かされている。また②の性質のおかげで、どんな由来であってもDNAの塩基配列にほぼ確実に違いがないことが保証されているし、微小な変異のおかげで個体差や血縁の解析が可能になる。最後に③の性質のおかげで、生理や知覚といった、直接観察困難な表現型の研究にもつなげることができる (Hayakawa, 2015)。ゲノムDNAにおける微小な変異の継承と蓄積が、表現型の自然選択につながり、生物進化の原動力となっていることも忘れてはいけな

2. 非侵襲サンプルによる野外DNAの分析

DNAを用いた野生霊長類の分子生態学研究の進歩は、分子生物学の技術革新とともにある。そもそもジェームズ・ワトソンとフランシス・クリックによって、DNAが遺伝現象の分子基盤であると提唱されたのが1953年のことである (Watson & Crick, 1953)。ゲノム研究自体がまだ半世紀あまりの歴史しかないのだ。野生霊長類の分子生態学研究に積極的にDNAが用いられはじめたのも1990年代に入ってからである。大型類人猿の野外調査地で採取された体毛、ワッジ (水分とそこに溶けだした栄養を吸ってから捨てられた繊維質の多い植物の残渣のこと)、尿などからミトコンドリアDNAの塩基配列を決定し、家系や多様性の解析に用いることが可能であるということがまず示された (Sugiyama *et al.*, 1993; Morin *et al.*, 1994; Hashimoto *et al.*, 1996; Hayakawa & Takenaka, 1999)。その後、核ゲノムに散在する変異性の高い反復配列 (microsatellite: マイクロサテライト) を用いて、血縁解析をおこなうことも可能となった。特にチンパンジーやマカクのような乱婚型の社会において、直接観察ではわからない父子関係の解明は、霊長類をはじめとする社会性動物の繁殖戦略の理解に大きく貢献している (Constable *et al.*, 2001; Vigilant *et al.*, 2001; Langergraber *et al.*, 2007; Inoue *et al.*, 2008; Inoue & Takenaka, 2008)。

分子生態学の基本は非侵襲サンプリングだ。実験動物では、潤沢なゲノムDNAを得るため、採血や生検を介して、血液中のリンパ球や上皮などの組織からDNAを抽出できる (Hofstetter *et al.*, 1997)。しかし野生動物から採血や生検をするためには、対象個体の捕獲が必要だ。特に大型動物において実施が困難だけでなく、対象個体や集団を高ストレス状態に置くことにもなり、本来知りたいはずの行動や生態への人為的影響や、動物福祉の観点から、できる限り避けたい。幸い、ゲノムDNAはあらゆる細胞に含まれているので、個体から自然脱落したものを利用することが可能だ (Eggert *et al.*, 2005)。先に述べたように、体毛の上皮細胞 (特に

毛根部)、ワッジなどの食べかすに付着した口腔細胞、排泄物 (尿や糞便) に含まれている膀胱や腸壁の細胞をDNAサンプルとして用いることができる。体毛はネストに残っているし、食べかすや排泄物は直接行動観察が可能なら、個体を特定した上でサンプリングができる。ときに、個体間の闘争などによって散乱した血液や組織片、経血、性行動によって残された精液を拾うこともできる (Hayakawa, 2008)。これらは個体への侵襲性がないという意味で、非侵襲サンプル (noninvasive sample) と呼ばれている。

非侵襲サンプルの欠点として、血液や上皮組織にくらべて得られるDNAが量的にも質的にも悪いという点がある (Morin *et al.*, 2001)。体毛や排泄物などの脱落物から得られるDNA量はとても少ない。また、糞便に含まれていると考えられている細胞は、消化管上皮のターンオーバーによって生じた腸壁細胞の老廃物であり、DNAの分解がはじまっていると考えられている。そのため、糞から抽出したゲノムDNAは断片化しており、塩基配列情報を十分に長く得ることができない。そうして、非侵襲サンプルをいかに効率よく野外で保存し、DNAを抽出して、分析に適用するかという手法が多くの研究で検討、開発されてきた。

DNAを構成するひとつひとつのヌクレオチドはリン酸を介した共有結合 (ホスホジエステル結合: phosphodiester bond) によってつながっている。ホスホジエステル結合は、生体サンプルには普遍的に存在するDNA分解酵素 (deoxyribonuclease, DNase) の存在下で、水分子による分解 (加水分解) を受けやすい。そのため、非侵襲サンプルに由来するDNAは分解が進んでいると考えられている (Frantzen *et al.*, 1998; Taberlet *et al.*, 1999)。酵素には温度依存性があるため、DNA分解酵素のはたらきにくい低温下、できれば水分子の影響を完全に抑えられる凍結状態でDNAサンプルを保存するのが好ましい。しかしながらフィールド調査地には冷蔵・冷凍設備がないのが通例である。そこでシリカゲルによってDNA分解の原因になる水分子を飛ばして乾燥保存を徹底したり、エタノールによって試料を脱水したり、あるいは加水分解酵素を阻害する界面活性剤や緩衝液中に保存したりするなどの手法が考案、実践されている。さらに、毛ではDNA抽出効率に影響するケラチンや、DNAから特定遺伝子領域を増幅するPCR (polymerase chain reaction: ポリメラーゼ連鎖反応) を阻害するメラニン、糞でもPCRを阻害する胆汁酸やビリルビンといった、DNA分析に影響する阻害物を多くの非侵襲サンプルが含んでいる。そのため、混合物を除去するDNA抽出を工夫することで、分析効率が格段にあがる (Wilson *et al.*, 1995; Zhang *et al.*, 2006)。PCRがDNA分析の基本であるが、DNA分析の基本であるが、量、質ともに悪いDNAを鋳型としたPCRはエラーを

引き起こす。繰り返し PCR をおこない、目的に応じて統計的に有意に信頼できる PCR 回数が考案されている (Morin et al., 2001)。こうした非侵襲サンプル利用の検討については井上 (2015) の総説に詳しい。

3. 次世代シーケンシング (NGS) の登場

野生動物を対象とした DNA 研究では、従来、PCR 法、電気泳動法、そしてサンガー法 (Sanger sequencing : サンガーシーケンシング) の 3 つが基礎となっていた。生物のゲノム DNA のサイズ (全体の塩基の長さ) は万から億というオーダーであり、一度に解析することはとてもできない。そこで知りたい遺伝子領域 (通常、数百~数千塩基長) の DNA のみを複製、増幅する技術として PCR は開発された (Mullis & Faloona, 1985)。電気泳動は PCR 産物 (アンプリコン: amplicon) をアガロースゲルなどに流し込み、電圧をかけることで分離する技術である。電気泳動で分離した PCR 産物を蛍光色素などで染色し、可視化することで、PCR 増幅が成功したかどうか、PCR 産物がどれくらいの塩基長を持つかなどを判定することができる。たとえば Y 染色体にしかない遺伝子をターゲットとして、PCR 産物の有無により、DNA 試料の由来が Y 染色体を持つオスであるかどうかを判定できる (Jayasankar et al., 2008)。また塩基長に個体差があるマイクロサテライト領域を PCR 増幅し、個体間で比較することで親子関係を明らかにする血縁解析なども可能だ (Queller et al., 1993)。こうした PCR 産物の塩基長をベースとした解析はフラグメント解析 (fragment analysis) と呼ばれ、ゲル電気泳動の代わりに一塩基レベルまで精度よく自動分離できるキャピラリー電気泳動もよく用いられている (Andersen et al., 2003)。

PCR 産物を構成する A, T, G, C の塩基配列を決定する手法 (sequencing analysis : シーケンシング解析) としては、サンガー法が用いられる (Sanger et al., 1977)。サンガー法は特定の塩基について伸長不可能な塩基 (ターミネーター: terminator) を混ぜて PCR 産物を再 PCR し、さまざまな長さを持つ断片を作成する。それを電気泳動にかけると、長さの順に A, T, G, C の塩基配列を復元することができる。サンガー法が考案された当時は放射性標識したターミネーターを用いて解析されていたが、現在は簡便な蛍光標識で代用し、キャピラリー電気泳動で自動判定するのが主流だ (Swerdlow & Gesteland, 1990)。キャピラリー電気泳動によるフラグメント解析とシーケンシング解析は、Applied Biosystems 社 (現在、Thermo Fisher Scientific 社のブランドの一つ) の Genetic Analyzer シリーズとして自動機器化されている。

キャピラリー電気泳動ベースのサンガーシーケン

シングでは、一本のキャピラリーあたり数百塩基程度の配列を読むことが可能だ。遺伝子のエクソン (タンパク質をコードする一続きの領域) の塩基配列はだいたい数百塩基程度なので、個別のエクソン配列を決定する研究では、PCR とサンガー法の組み合わせは非常に有効な手法となっている。しかし、哺乳類ゲノム全体のサイズは 30 億塩基である。その 1.5% に、数万個ものタンパク質をコードする遺伝子が刻まれている (International Human Genome Sequencing Consortium, 2001)。残りの 98.5% も、遺伝子発現の制御領域やイントロン、反復配列など、ゲノムを理解する上で不可欠な領域が多く含まれている。生物をゲノムレベルで理解するためには、この 30 億個の塩基配列を完全解読する必要がある。そうすると、ヒトゲノム全体を一回読むためには、単純計算で数千万回のキャピラリーでのサンガーシーケンシングが必要だ。さらに一回読むだけでは信頼性が低いので、同じ領域を何十回も読む必要がある (この回数のことを深度 depth と呼ぶ)。2003 年に完了したヒトゲノム研究の国際コンソーシアムでは、このとてつもない回数のサンガーシーケンシングを実施して、ヒトゲノムの全長配列を高精度で決定した。2800 人を超える科学者が参加し、27 億ドルの予算を投じて、13 年の年月を費したと報じられている (Wadman, 2008)。

ゲノムプロジェクトでは、ヒト以外の多種多様な生物でも、全ゲノム配列がサンガーシーケンシングにより解読された。マウス (*Mus musculus*)、シロイヌナズナ (*Arabidopsis thaliana*)、大腸菌 (*Escherichia coli*) といった実験室で欠かせないモデル生物はもちろん、非ヒト霊長類でもチンパンジーやアカゲザル (*Macaca mulatta*) のゲノムが解読された (Blattner et al., 1997; Arabidopsis Genome Initiative, 2000; Mouse Genome Sequencing Consortium, 2002; Chimpanzee Sequencing and Analysis Consortium, 2005; Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007)。ゲノム研究の発展は同時に、技術革新ももたらした。それがサンガーシーケンシングにとらわれない NGS (next-generation sequencing : 次世代シーケンシング) と呼ばれる技術である。先端技術を活用した新しいシーケンシング手法の総称であり、2007 年ごろに実用化された。パイロシーケンシング、ブリッジ PCR、シーケンシング・バイ・シンセシス、半導体シーケンシング、一分子シーケンシングなど、さまざまな手法が含まれている (Heather & Chain, 2016; Shendure et al., 2017)。

原理はさまざまだが、これらの NGS 技術は共通して、サンガー法で使われていた電気泳動という制約を完全に取り払った。ひとつは空間的なブレイクスルーで、サンガー法ではキャピラリーの一本一本で個別に配列

を決定していたが、NGS では読みたい DNA 分子を超極小スケールの基盤上に数万から数千万本という規模で配し、それぞれを読むことができるようになった。もうひとつは時間的なブレイクスルーで、サンガー法ではあらかじめターミネーターを付加した PCR 産物を用意する必要があったが、NGS では DNA 分子をそのまま基盤に添加してから、シークエンシング反応をリアルタイムに測定する。合成しながらのシークエンシングという意味で、シークエンシング・バイ・シンセシス法 (SBS: sequencing by synthesis) と呼ばれている (Hyman, 1988; Ronaghi *et al.*, 1998)。こうしたブレイクスルーによって、ギガ (10 億) からテラ (1 兆) 塩基規模の配列を一度に得ることを可能とした。つまり、哺乳類ゲノムを構成する 30 億もの塩基配列を、NGS はたった一回運用することで読むことができるので、鋳型とする DNA サンプル量も節約できる。コストも格段に下がった。次世代シーケンサーの開発会社のひとつである Illumina 社は、2014 年に HiSeq X シリーズという次世代シーケンサーをリリースし、これを用いると 1000 ドルでヒトゲノム解読が可能であると宣言した。27 億ドルを必要とした 20 年前から、1000 ドルゲノム時代への突入は、半導体などの他のテクノロジーの進歩のスピード (ムーアの法則と呼ばれる) からはるかに逸脱した速さだといわれている (Hayden, 2014)。

4. 霊長類遺伝学から霊長類ゲノミクスへ

2003 年のヒトゲノム解読を筆頭として、さまざまなモデル生物の全ゲノム配列が解読された当時、ポストゲノム時代に突入したと称された (Medini *et al.*, 2008; Eisenberg *et al.*, 2000)。DNA は RNA (ribonucleic acid: リボ核酸) を介してタンパク質に変換されることで遺伝子という機能を持つ。ヒトゲノム中にコードされているタンパク質は数万種類あると推定されているが、その 1 つ 1 つの個別研究を統合するのは容易なアプローチではない。しかし、その数万種類のタンパク質の情報を「わずか」3 ギガの ATGC の文字列情報 (1 枚の DVD ディスクに収められる程度の情報である) として集約することができれば、タンパク質が織りなす生命活動の全体像をとらえられると期待された。こういう見方から、DNA からタンパク質への変換の仕組みは「生命のセントラルドグマ (中心教義)」であり、ゲノムは「生命の設計図」とか「生命の青写真」などとも比喻された。

ポストゲノムのアプローチは霊長類の遺伝学研究の手法も大きく変えた。たとえば筆者らは、ゲノムプロジェクトによって解読されたチンパンジーの全ゲノム配列 (Chimpanzee Sequencing and Analysis Consortium, 2005) をリファレンス配列として用い、複数個体のチ

ンパンジーの苦味受容体遺伝子をサンガー法で塩基配列を決定した。その結果、チンパンジー亜種間で苦味受容体遺伝子が自然選択をとまって分化していることを明らかにできた (Hayakawa *et al.*, 2012)。PCR で特定の遺伝子領域を増幅するためにはその領域をはさむ部分配列 (プライマー) を知っておく必要があるが、全ゲノムリファレンス配列があるおかげで、そうしたプライマーを簡便に作ることが可能だった。全ゲノム配列のわかっていない生物種ではこうはいかなかっただろう。ゲノムの読まれた生物において、遺伝子の個体差の研究 (遺伝学) が飛躍的に進んだというのが、ポストゲノムのアプローチがもたらした結果のひとつといえる。

一方で、ポストゲノム研究においてもまだ、リファレンス配列のない生物の研究は難しいということ、そして研究の基盤がスループットの低いサンガーシークエンシングであるという二点の制約が残っていた。NGS の登場はこれらの制約を完全に取り払った。NGS は領域特異的なプライマーを必ずしも要求としないので、リファレンス配列がない生物であっても適用可能であるし、得られるデータ量もサンガーシークエンシングに比べ圧倒的に多い。2000 年代に発表された霊長類ゲノムはヒト、チンパンジー、アカゲザルのみであったが、NGS が一般化した 2010 年代には、NGS によって解読された非ヒト霊長類ゲノムの発表が加速化し、2011 年から 2012 年の間に、スマトラオランウータン (*Pongo abelii*), ニシゴリラ (*Gorilla gorilla*), ボノボ (*Pan paniscus*), フィリピンメガネザル (*Carlito syrichta*), オオガラゴ (*Otolemur garnettii*), ネズミキツネザル (*Microcebus murinus*), アイアイ (*Daubentonia madagascariensis*) の 7 種類のゲノムが次々に発表された (Lindblad-Toh *et al.*, 2011; Locke *et al.*, 2011; Perry *et al.*, 2012; Prüfer *et al.*, 2012; Scally *et al.*, 2012)。現在では、全ゲノム配列をデータベース化した Web サイトである Ensembl (<https://asia.ensembl.org>) を見ると、24 種類もの霊長類ゲノムが登録されている (Ensembl Release 91: 2017 年 12 月版)。筆者らも、こうした NGS で解読されたデータを含む 15 種類の霊長類の全ゲノム配列を比較して、苦味受容体遺伝子のレパートリーが食性進化と対応して変化していることを明らかにした (Hayakawa *et al.*, 2014)。PCR やサンガーシークエンシングを基盤としていた霊長類の遺伝学研究は、NGS によるゲノムレベル解析を基盤としたゲノム研究 (genomics: ゲノミクス) へと大きく転換した (Rogers & Gibbs, 2014)。

「生命のセントラルドグマ」は RNA を介した DNA からタンパク質への変換機構であるとされたが、実際にはタンパク質はゲノムのわずか 1.5% の領域にしかコードされていない。残りの膨大な非コード領域にも、タンパク質の発現を制御する配列や、単体で機能を持つ

RNA 配列など、重要な情報が刻まれている。ヒトゲノムやマウスゲノムを筆頭に次々に解読された 29 種の有胎盤哺乳類ゲノムを比較すると、ゲノム中の約 4.2% が保存されていることがわかった (Lindblad-Toh *et al.*, 2011)。種を超えて進化的に保存されているのならば生命として不可欠な情報だと考えられる。この数値はタンパク質をコードする 1.5% を上回るので、タンパク質のすべてがわかれば様々なことがわかるというポストゲノムという言葉に込められた当初の期待ほど、生命活動のメカニズムは解明されていない。

5. 野生霊長類ゲノム研究への NGS の適用

一度にギガスケールの塩基配列を決定でき、リファレンス配列を必要とせず、鋳型とする DNA も少量で済むといった NGS の特徴は、非侵襲サンプルに基づく野外研究でも非常に実用性が高い。たとえば筆者の経験では、野外で集めた野生霊長類の糞サンプルのスワブ (ぬぐい液) から、数マイクログラム程度の DNA が採取できる。一回の PCR で百ナノグラム程度の DNA を PCR の鋳型として使うとすると、よくて百回程度しかサンガーシーケンシングができない。つまりどんなに効率よく実験ができて、ひとつの糞サンプルの分析で数万塩基長の配列しか決定できない。とても哺乳類が持つ 30 億塩基のゲノムサイズには太刀打ちができない。しかし、NGS はこうした少量の DNA でも数十億塩基の配列、すなわち哺乳類ゲノムひとつ分の配列を理論的には決定できる。リファレンス配列を必要としないので、どんな野生動物にも適用できる。分子生態学において NGS を適用することは非常に高い可能性を秘めている。

実際には、一塩基あたりの単価は各段に安いといっても、一回あたりの NGS の運用 (ラン) にかかる経費が高いため、数ギガ規模のデータ量を必要としない研究ではサンガーシーケンスの方が実用的だ。また、情報解析に必要なソフトウェアの開発が分子生態学研究向けに進んでいない理由などから、野生動物のゲノム解析の分野に普及しているとはいえない。しかし、NGS 技術を野生動物のゲノム研究に応用するアイデアはすでにいくつか提示されている。

2000 年代後半に NGS が普及して、野生動物研究への応用として最初に適用されたのが、血縁解析を目的とした非モデル生物におけるマイクロサテライトマーカーの開発だ。マイクロサテライト (STR: short tandem repeat または SSR: simple sequence repeat ともいう) とは、ゲノム中に散在する 1 塩基～6 塩基を単位とした繰り返し配列のことで、たとえば CACACACA (CA の 4 回繰り返し) とか、AATAATAAT (AAT の 3 回繰り返し) というような配列をいう。ヒトゲノムでは 3% がマイク

ロサテライトで占められている (International Human Genome Sequencing Consortium, 2001)。タンパク質をコードするエクソン領域 (1.5%) の 2 倍の割合だと考えると、いかにマイクロサテライトがゲノム中に普遍的に散在するものと理解できる。繰り返し配列は相補鎖とのずれ (slippage) を簡単に引き起こし、ずれが生じたまま DNA 複製が起こると、相同染色体間で異なる繰り返し回数の配列が創出される。これが、マイクロサテライトがゲノム中に出現しやすく、かつ繰り返し数の多型が非常に多い理由だと考えられている。ゲノム中に散在し、多型が生じやすいことから、マイクロサテライトは NGS の登場以前から、血縁解析のスタンダードとして利用されてきた (Queller *et al.*, 1993) が、マイクロサテライト領域を確実に PCR するためのプライマーの設計は決して容易ではなかった。しかし、全ゲノム配列が未知である生物種であっても、NGS でランダムにゲノム配列を大量に読むことで、よければ数パーセントの確率でマイクロサテライトを発見できる。特段のリファレンス配列なしに NGS で配列を特定することを *de novo* シーケンシングという。

NGS を用いた最初のマイクロサテライト開発研究のひとつとして、アオヤマガモ (*Hymenolaimus malacorhynchos*) での研究が報告された (Abdelkrim *et al.*, 2009)。アオヤマガモのゲノム DNA を断片化し、最初に普及した NGS である Genome Sequencer FLX (GS-FLX: Roche 社) を用いて 17,215 リードの配列を得たところ、231 リードがマイクロサテライトであり、最終的に 13 個のマイクロサテライト領域を多型サイトとして実用化できたとされている。また、この論文が掲載された同誌には、こうした NGS によるマイクロサテライト開発がさまざまな分類群の生物 (真菌, 昆虫, 線虫) に適用可能であるということも報告されている (Santana *et al.*, 2009)。10 か所以上のマイクロサテライトマーカーが開発できると、血縁解析や遺伝的多様性の評価に有用なため、2009 年以降、さまざまな非モデル生物で開発が試みられている (Adenyo *et al.*, 2012; Akomo-Okoue *et al.*, 2013; Schoebel *et al.*, 2013)。霊長類はすでにさまざまな種で全ゲノム配列が明らかにされているおり、マイクロサテライト領域の候補も容易に計算機上で検索できるので、霊長類を対象に *de novo* シーケンシングでマイクロサテライトマーカーを開発する場面はあまりないが、NGS を野生の非モデル生物研究に活かすパイオニア研究として、言及する価値があるだろう。

6. 野生霊長類におけるゲノムワイド SNP 解析

NGS のスループットは日進月歩で増大しており、マイクロサテライトではなく一塩基多型 (SNP: single nucleotide polymorphism) をベースとした解析でも、低

コストで血縁解析ができるということが示されはじめています。NGSで網羅的に調べた配列を、リファレンス配列にマッピングすることで、容易にSNPを検出することができます。これをリファレンススペースの全ゲノム解析という意味でリシーケンシング (resequencing) と呼ぶ。2015年には、ヒトゲノムのリファレンス配列をベースに、世界中の1000人を対象としたリシーケンシング研究、通称1000人ゲノムプロジェクトの結果が報告されている (1000 Genomes Project Consortium, 2015)。そして、今では多くのパーソナルゲノム研究が世界中で進んでいる。

野生霊長類の研究では、ノートルダム大学が指揮している Amboseli Baboon Research Project において、長期個体群調査がされている野生のヒヒ (キイロヒヒ *Papio cynocephalus* とアヌビスヒヒ *P. anubis* の自然交雑域) での研究がある (Snyder-Mackler *et al.*, 2016)。あらかじめ家系を把握している個体識別された62個体の野生のキイロヒヒの糞DNAから、後述するキャプチャー法を組み合わせてホストゲノムDNAをできる限り濃縮して、低深度で全ゲノム配列をリシーケンシングした。そして塩基多型に基づいて遺伝関係を再構築した。その結果、非常に高い精度で家系図を再現できたと報告している。高度な実験技術やそれなりの経費がかかるため、まだこうしたアプローチが野生霊長類研究の現場に普及するには年月を要するだろうが、いずれマイクロサテライト解析やサンガー法による血縁解析に並ぶスタンダードな手法となることを期待させる。

NGSによるゲノムワイドなSNP解析は、血縁解析以外にもさまざまな研究の関心に活用できる。ひとつは集団や種の遺伝的多様性の評価、もうひとつは表現型解明である。本稿の冒頭で、世代あたりに一定の割合 (突然変異率) でゲノムの塩基配列に微小な変異が入り、その蓄積が個体差や種差につながっていると述べた。個体差があるということは、集団の遺伝的多様性が高いということだ。集団遺伝学の分野では、こうした塩基変異と集団多様性との関係について理論が確立している (Nei & Kumar, 2000)。すなわち、哺乳類のような二倍体のゲノムにおいて、個体間における塩基あたりの塩基置換の割合を塩基多様度 (π , nucleotide diversity), 一世代あたりの突然変異率を μ としたときに、その集団の有効サイズ N_e (繁殖に寄与している個体数) は中立下 (自然選択をとみなわない) では

$$\pi = 4N_e\mu$$

に従うとされている。この式に基づいて、遺伝的多様性の評価や機能遺伝子における自然選択の検証などさまざまな分子進化研究が実施されている。たとえば筆者らは、チンパンジーの苦味受容体遺伝子がどのよう

な自然選択圧下にあるかを検証するために、SNPデータから塩基多様度を計算した (Hayakawa *et al.*, 2012)。従来は集団のSNPデータを、領域特異的にPCRをし、サンガーシーケンシングを多数おこなうことで得ていたが、NGSを用いることで、容易に、かつゲノムワイドにSNPデータを得ることができる。さらに π 値は、二倍体ゲノムではゲノムワイドなヘテロ接合度 (相同染色体どうしで塩基が異なる割合) で置き換えることができる。このモデルを利用して、一個体の全ゲノムのヘテロ接合性のパターンを解析すると、その個体が属する集団の有効サイズ (N_e) がどのような変遷の歴史を歩んできたかを推定できる。現在観察できる遺伝的多様性をもとに祖先状態を推定するという意味で、合祖理論 (coalescent theory) と呼ばれている。

NGSによるSNPデータに合祖理論を適用するモデルとしてPSMC (pairwise sequentially Markovian coalescent) が2011年に開発された (Li & Durbin, 2011)。PSMCをゲノムワイドSNPデータに適用すると、定量的に、その種が何百万年前にどれくらいの集団サイズであったかということを推定することができる。これをヒトゲノムデータに適用すると、ヒトは *Homo* 属が出現して以降、1万から2万個体程度のレンジで有効集団サイズを変動させて、出アフリカ後、アジアやヨーロッパ集団で創始者効果によるボトルネックが生じたということが、見事に再現される。化石でしかわからなかった過去をゲノムから推定できるようになったのだ。PSMCは現在、スタンダードな解析となり、どんな動物でも全ゲノムシーケンシングをおこなったときは必ず実施するようになっている。合祖理論の考え方はヒト科全種にも適用されており、驚異的なことにヒト、チンパンジー、ゴリラ、オランウータンの共通祖先がどんな集団構造を持っていたかというはるか昔の歴史まで推定されている (Prado-Martinez *et al.*, 2013)。

生物種の遺伝的多様性の保全を考える上で、集団の個体数を維持することが重要であるということを私たちは直観的に知っている。この概念は、先に述べた塩基多様性と集団サイズが比例関係にあるという式によって、定量化することができる。NGSによるゲノムワイドSNP解析は、保全生物学の領域でも新たなツールとして活用されていくだろう (Arandjelovic & Vigilant, 2018)。

7. 糞DNA中のホストゲノムのキャプチャーシーケンシング

野生集団の遺伝的多様性をゲノムワイドで解析するためには、やはり非侵襲サンプルを用いる必要があるという制約が出てくる。非侵襲サンプルは微量であるだけでなく、対象動物 (ホスト) 以外のDNAも混入し

てくる。たとえば非侵襲サンプルとして最もよく使われる糞からは、ホストの腸壁細胞の老廃物由来の DNA が抽出できるが、同時に消化後の食べ物の残渣や、共生している腸内細菌の DNA も抽出されてしまう。糞の状態によってばらつきはあるが、糞 DNA の約 99% はこれら非ホスト由来のものである。この DNA をそのまま NGS にかけて、99% は無駄になってしまう。そういった問題を解決する方法のひとつがキャプチャーシーケンシング (capture sequencing) である (Ng *et al.*, 2009)。その名の通り、事前に目的の DNA のみをつかまえて (キャプチャーして) 濃縮し、それから NGS にかけてという技術である。DNA の塩基配列は、A に対しては T に、G に対しては C に、強固な水素結合を形成する。その結果、相補的な配列と二重螺旋構造を形成できる。この仕組みを利用して、キャプチャーしたい配列と相補関係にある短い DNA (または RNA) プローブを人工的に合成し、DNA 溶液中で相補結合を形成させる (ハイブリダイゼーションという)。プローブにはあらかじめ、磁気ビーズを付加しておく。磁力をかけてハイブリダイゼーションした DNA 分子だけを分離することで、濃縮ができる。さいわい、霊長類ではリファレンス配列がわかっているので、リファレンス配列をベースにプローブを合成することで、キャプチャーが容易にできる。糞サンプルからのキャプチャーシーケンシングのアイデアは 2010 年に提示され、飼育下のチンパンジーの糞から合計 150 万塩基の核ゲノムおよびミトコンドリアゲノム全長の配列決定に成功している (Perry *et al.*, 2010)。

プローブの人工合成をするのは非常にコストがかかるため、ふたつの代替法が実践されている。ひとつは先に述べた Amboseli Baboon Research Project のヒヒに用いられた手法だ (Snyder-Mackler *et al.*, 2016)。血液由来などのクオリティの高いヒヒゲノム DNA を用意し、それを適切なサイズに断片化して、各断片の末端に磁気ビーズを付加して、プローブにした。人工合成の必要がないのでコストは削減できるし、飼育下などでクオリティの高い DNA が入手できればどんな生物でも適用できるので、今後普及していく手法であろう。

もうひとつはヒトのプローブを用いることである。ヒトは言うまでもなく、もっとも生物学研究がされている霊長類であり、キャプチャーシーケンシングのためのプローブも製品として市販されている。カスタムで合成する必要がないのでコストが低い。非ヒト霊長類はヒトと進化的に近縁なため、塩基差異はそこまで大きくない。ヒトで作成されたプローブで十分にハイブリダイゼーションできる。

ヒトのキャプチャーシーケンシングでよく実施されるのがエクソーム解析 (exome sequencing) だ。ヒトゲノムで、実際にタンパク質をコードしているエクソ

ンの領域は全体の 1.5% しかない (International Human Genome Sequencing Consortium, 2001)。医学などの関心でタンパク質変異に関心がある場合は、全ゲノムを読まなくてもこの 1.5% のエクソンを対象とすればよい。そこでエクソン特異的なキャプチャープローブが製品化されており、エクソン全体という意味でエクソームと名付けられている。実際に、チンパンジーやアカゲザルでヒトのエクソームプローブが適用可能であることが示されており (Vallender, 2011)、2017 年には 18 個体の野生のチンパンジーから採取された糞からエクソーム解析に成功したと報告された (Hernandez-Rodriguez *et al.*, 2017)。その 18 個体のうち、9 個体はヒガシチンパンジー (*P. t. schweinfurthii*)、残りの 9 個体はチュオウチンパンジー (*P. t. troglodytes*) であるが、エクソームワイドの SNP データで十分に亜種が識別できることが確認されている。タンパク質をコードする領域の変異は、そのまま表現型に直結するので、こうしたエクソーム解析の技術は、血縁解析や遺伝的多様性の評価のみならず、野生集団における遺伝子型—表現型解析に向けて大きく踏み込んでいけるだろう。

8. NGS による環境 DNA 解析

NGS の登場は、野生霊長類のゲノム研究だけでなく、環境 DNA (environmental DNA) の研究にも大きな革新をもたらした (Bohmann *et al.*, 2014)。実用性の高さや実践例の豊富さという点では、環境 DNA 研究の方が躍進しているといえる。ここでいう環境 DNA とは、採取した遺伝試料に含まれている対象動物 (ホスト) 以外の DNA のことを指す。糞サンプルであれば、ホストゲノムの割合はわずか 1% で、残りの 99% は消化後の食べ物の残渣や、共生している腸内細菌から構成されていると考えられる (Perry *et al.*, 2010; Human Microbiome Project Consortium, 2012)。これら非ホスト由来 DNA はホスト DNA を解析する上では不要な障壁で、キャプチャーなどの工夫を余儀なくされた。しかし、それを逆手に取ると、NGS で網羅的に非ホスト DNA を解析することで、どんな食べ物を食べているか、どんな腸内細菌が共生しているかといった生態情報を知ることにつながることもいえる。

そもそも糞分析は動物生態学で欠かせない手法のひとつだ。とりわけ人なれしておらず、行動を直接観察できない野生動物を対象とするとき、糞を拾って洗い、含まれている種子や無脊椎動物の断片を調べることで、その動物の食性をあきらかにできる。また、新鮮な糞を集め、培地に接種し、適切な条件で培養、分離することで、どんな機能をもった腸内細菌が共生しているかも調べることもできる。環境 DNA 解析はこうしたマクロな糞分析に代わる手法といえるだろう。

環境 DNA を対象とした研究は、ホストゲノム研究に対し、高次 (meta-) のゲノムという意味で、メタゲノム (metagenome) 研究 (メタゲノミクス: metagenomics) と呼ばれる。特に腸内細菌などの共生微生物の遺伝子解析に使われる言葉で、遺伝子の総体をゲノムと定義したように、共生微生物が持つそれぞれの遺伝子がまるで種を超えて一体となってホストの生命活動に影響するように見えることからそうのように名付けられた。糞の 99% は環境由来 DNA ということもあり、環境 DNA はホストよりも非常に解析しやすく、霊長類を含む多くの野生動物の食物レパートリーや腸内細菌フローラの網羅的同定を目的として、多くの研究が NGS を利用している (Ley *et al.*, 2008; Muegge *et al.*, 2011; Human Microbiome Project Consortium, 2012; Srivathsan *et al.*, 2015)。糞以外でも、例えば口腔スワブ試料に含まれている口腔細菌や、土壌や水圏に含まれている DNA からその生態系に属する生物種の同定などで応用されている (Bohmann *et al.*, 2014)。

9. メタバーコーディング解析とメタゲノミクス解析

NGS による環境 DNA の研究は 2 つの手法に分けられる。ひとつはメタバーコーディング解析 (metabarcoding analysis; アンプリコン解析とも)、もうひとつはメタゲノミクス解析 (ショットガン解析とも) である。メタバーコーディング解析は、対象とする生物分類群が共通して持つ遺伝子領域 (バーコード配列) を PCR で一度増幅し、PCR 産物を NGS で網羅的に決定する手法である。NGS が普及する以前は、PCR 産物をおおよそ一分子ずつ大腸菌に組み込み、多数のコロニーからもう一度 PCR することでひとつひとつシークエンシングするという手法 (クローニング法) が用いられていた。サンガーシークエンシングでひとつひとつ読まなくては行けないため、時間的にも経費的にも、NGS の方がリーズナブルである。

バーコード配列には、種間でよく保存されている領域 (保存領域) と、種間で変異に富む領域 (可変領域) の両方を含んでいる必要がある。保存領域は遺伝子機能としても重要で自然選択が生じている場所、可変領域は機能的にあまり重要ではなく中立的に変異が蓄積している場所とも読みかえることができる。保存領域に PCR プライマーを設計してどんな種が対象でも増幅がかかるようにし、増幅された可変領域で分類群を識別するという理屈である。動物ではミトコンドリアゲノムの COI 遺伝子、植物では葉緑体ゲノムの *matK* や *rbcl* 遺伝子、真菌では核ゲノムの ITS 領域などがよく用いられる (Folmer *et al.*, 1994; CBOL Plant Working Group, 2009; Pompanon *et al.*, 2012; Toju *et al.*, 2016)。霊長類は動物も植物も真菌も採食するので、メタバーコーディング解析によって食物品目を調べるためには、そ

れぞれの高次分類群に合わせて PCR をする必要がある。

メタバーコーディング解析は、PCR 産物のみの配列決定で済むため実験にかかる経費も安く、得られたバーコード配列がそのまま含まれている生物種情報を反映するため解釈も難しくない。さらに SBS 法ベースの Illumina 社のシークエンサーでは、あらかじめシークエンスアダプターを付加したプライマーで PCR をすることで、精製した PCR 産物をそのままシークエンサーにセットすることができるため、特別な実験手技も必要としない。そうした分子生態学との親和性の高さから、アンプリコン解析はもっとも野生の霊長類を対象として NGS が実践されている手法だろう。シークエンス結果を集約 (アセンブル) し、バーコード配列が由来した分類群の同定や、配列数の豊富度の解析に必要なソフトウェアも、初学者にもやさしい形でパッケージ化されて無償で配布されており、Qiime や Claident などが利用可能である (Caporaso *et al.*, 2010; Tanabe & Toju, 2013)。

一方、アンプリコン解析にも簡便であるがゆえのデメリットは多数ある。アンプリコン解析では、バーコード配列をはさむ保存領域の配列をユニバーサルプライマーとして用いて PCR 増幅するが、生物種によってはまれにユニバーサルプライマーから大きく外れた塩基変異を持つことがある。そうした種は当然、アンプリコン解析では検出されないため、結果の解釈を誤ることがある。また、解析可能なバーコード配列の長さは数百塩基程度で、変異が少ない分類群では種を識別できず、よくて科レベルまでしかわからない。共生生物 (腸内細菌など) がどんな機能の遺伝子を持っていて、ホストの生命活動に貢献しているかについての情報は、バーコード配列からは全くわからない。

メタバーコーディング解析の弱点をカバーするには、メタゲノミクス解析という似て非なる解析を実施する必要がある。ショットガン解析 (shotgun metagenomic sequencing) と呼ばれ、環境サンプル (糞便、土壌、水など) から抽出した DNA を、PCR を介することなく、断片化 (ショットガン) して、NGS の圧倒的なスループットに任せて完全網羅的に配列決定することをいう。PCR によるバイアスがないので、特定の生物種を見落とすこともないし、塩基長の制限もないので種レベルでの同定も理論上は可能である。さらに機能をもった遺伝子を調べることができるので、共生している生物がどんな役割を持っているかも明らかにすることができる。しかし、想像に難くないように、メタゲノミクス解析は、そのサンプルに含まれている生物種のすべてのゲノムが対象になるので、得られる情報量は膨大である。哺乳類の糞であれば 1% がホストに由来するので、糞 DNA のメタゲノミクス解析の結果の 1% には、30 億塩基から構成されるホストゲノムが混ざってくる (Human

Microbiome Project Consortium, 2012)。データが得られた後の情報解析が非常に困難だけでなく、解析に耐えうる膨大なシーケンス量を得る必要があるため、経費もかかる。

実際にドゥクラングール (*Pygathrix nemaeus*) の糞に含まれる採食植物 DNA を対象として、メタバーコーディング解析とメタゲノミクス解析を比較した研究では、予想どおりの結果が得られている (Srivathsan *et al.*, 2015)。すなわち、メタバーコーディング解析では、配列が短すぎたために、植物を科より細かいレベルまで判別することが難しかった一方で、メタゲノミクス解析では、シーケンス量に限界があったために、まれな植物種は検出できなかったと報告されている。対象のサンプル数が大量であったり、細かい種情報よりもサンプル間の種構成の差異の比較を目的としたりする場合にメタバーコーディング解析、細かい分類群を知る必要があったり、共生生物の遺伝子の機能まで知りたい場合には、メタゲノミクス解析を選択するというような検討が、環境 DNA 研究を実践するにあたり必要である。

10. 腸内細菌叢解析のインパクト

NGS は腸内細菌の研究にも大きなパラダイムシフトを与えた。たとえばヒトの大腸には、およそ千種、細胞数では兆を超えるとされる細菌が共生しており、食べ物の消化や代謝に貢献している。細菌が草のように腸管に生い茂っているように想像されることから、叢 (くさむら) にたとえて腸内細菌叢 (enterobacterial flora) と呼ばれている (Eckburg *et al.*, 2005)。腸内細菌叢の研究は、食生活、生理、病気などの理解に欠かせない。以前までの細菌の研究では単離と培養というプロセスが必要で、培養ができない細菌についての情報は手ばかりがほとんど得られなかった。シーケンシングの技術が進むと、メタバーコーディング解析の手法を用いれば、培養を介さずともどんな分類群の腸内細菌が糞便として排出されているかを明らかにできるというアイデアが提示された (Amann *et al.*, 1995)。そして NGS が普及し、腸内細菌解析の分野にも NGS が積極的に用いられるようになった。

細菌の系統解析には、16S rRNA (リボソーム RNA : ribosomal RNA) の配列が最もよく使われている (Weisburg *et al.*, 1991)。16S rRNA をコードする DNA 配列にも、保存領域と可変領域があるためである。さまざまな細菌種の 16S rRNA 配列を収集したデータベース (DeSantis *et al.*, 2006) も構築されている。糞 DNA の 16S rRNA をターゲットとしたメタバーコーディング解析によって腸管に共生する細菌の豊富さや系統分類などを即座に解析することができる。また 16S rRNA に限

定しないショットガン・メタゲノミクスの方法論も確立している (Sczyrba *et al.*, 2017)。

非ヒト霊長類においても、すでに多くの NGS ベースの腸内細菌叢研究が発表されており、さまざまな新しい知見が提示されている。NGS による腸内細菌解析が最初にもたらしたインパクトの高い概念がエンテロタイプ (enterotype) だ (Arumugam *et al.*, 2011)。ヒトの腸内細菌叢を調べると、国籍や性年齢などには大きく影響を受けずに、食習慣や投薬履歴によって腸内細菌叢のパターン構成が 3 種類に分類されることが明らかにされた。こうしたパターンをエンテロタイプとして定義している。もちろん性年齢などのさまざまなホストの特性の影響を受ける微小な構成は多数あるが、ホストの食環境によって形成される共生微生物の新たな表現型として話題を呼んだ。その後、同様の解析をチンパンジーの腸内細菌叢に適用すると、やはりエンテロタイプのようなパターンが検出され、ホスト-微生物共生関係のあらたな概念として注目された (Moeller *et al.*, 2012)。

他にも、菌叢パターンにはホストの系統によって制約があること (Sanders *et al.*, 2014; Moeller *et al.*, 2016a)、季節によって利用できる食物が変動するような環境に生息している霊長類では、腸内細菌叢も変動するという事実 (Amato *et al.*, 2015; Sun *et al.*, 2016; Smits *et al.*, 2017) や、前胃を持つ葉食性のコロブスは飼育下で消化管疾患にかり、腸内細菌叢にも影響が出ること (Amato *et al.*, 2016)、群れの中で親和的な関係にある個体間では腸内細菌叢も類似してくるというネットワークとの関係 (Tung *et al.*, 2015; Moeller *et al.*, 2016b; Perofsky *et al.*, 2017)、そして飼育下の霊長類は野生とは離れ、西洋食的になってしまうことから、菌構成もヒトのようになってしまうこと (Clayton *et al.*, 2016) など、さまざまな新しい概念が提示されている。野生動物のお腹の中を直接観察することはできないが、そこに存在している微生物生態系が、社会性動物である霊長類の個体間ネットワークに関係したり、飼育と野生でドラスティックに菌叢が変わってしまったりするのだ。微生物学の研究を超えて、霊長類生態学や保全生物学の観点からも欠かせない視点になるだろう。

共生細菌は腸管にとどまらず、上皮や口腔などでも生態系を構築している。これらのスワブを得ることができれば、腸内細菌同様に菌叢を解析できる。これらの微生物が構成するコミュニティ全体を、特に NGS の分野でマイクロバイオーム (microbiome) と呼ぶことが多い。ここでもゲノム、エクソーム、メタゲノムの名で使ったようなオーム (-ome) という言葉を使う。NGS でまるごと全体を解析したいという意図があつてのことである。糞であれ、スワブであれ、野生の霊長類のマイクロバイオームを NGS で解析するには、やは

り保存が問題になってくる。検出さえできればよいホストゲノムや、定量性を必要としない食物アイテム（多細胞生物であるため、DNA量が摂取量を反映する保証がまったくないため）とは異なり、単細胞生物であり、DNAのコピー数が細胞数をそのまま反映するマクロバイオーム解析では、種構成や場合によっては各種の豊富さ（PCRによって増える割合）なども分析対象になってくる。野外での保存中や実験室への運搬中に、特定の菌種が増殖したり、あるいは死滅したりすると、結果が大きく変わってしまう。とくに腸内細菌は嫌気性細菌が多くを占めるため、糞として体外に放出され空気に触れた瞬間に、菌の死滅がはじまってしまう。できる限り直接観察による速やかな収集が必要である。野外で利用できる収集法として、ホストゲノムの保存にも使われるような乾燥、エタノール、RNAlater、界面活性剤が提案されている。どれも冷凍保存で得られる結果を完全には再現しないが、どんな菌種が存在したかといったことはある程度再現できるということが、いくつかの検証研究で確認されている（Choo *et al.*, 2015; Flores *et al.*, 2015; Gorzelak *et al.*, 2015; Hale *et al.*, 2015; Song *et al.*, 2016）。各種の保存法の一長一短をあらかじめ把握し、異なる保存法によって得られたサンプルを直接比較はしないなどの注意を払って、野外でのサンプリングを実施すべきだろう。

11. 非侵襲サンプルを用いたメタ分析の展望

冷蔵庫や冷凍庫のないような調査地で、野生の動物を対象とする以上、質のよいDNAを収集、保存することは容易ではない。こうした微量で、質の悪いDNAであっても、NGSは分析を可能としている。逆に質の悪さは織り込み済みで、糞のDNAからホストゲノムだけでなく採食アイテムや腸内細菌を解析することは、野外研究の可能性を提示している。非侵襲サンプルはDNAだけでなく、ホルモンや、安定同位体の分析にも使用することができるので、分野を超えた連携によって分析手法を一つのサンプルに適用することは、野生霊長類の新しいメタ分析に繋がるだろう。

たとえばDNAから性判定をすることはできるが、DNAは身体のどの細胞から採取しても情報が変わらないため、個体内の生理的な変化まで知ることはできない。そこでホルモンなどの代謝物を同じサンプルから同時に調べることで、その試料に由来した個体がどちらの性別で、どんな生理状態にあるかを知ることができる。あるいは糞DNAを用いて、どんな種類の植物を食べたかを知ることができるが、植物のどの部位（葉か果実かなど）を食べたかという情報はDNAからはわからない。直接観察やカメラトラップなどの行動を見るという手法に取って変わるものではない。また、別

の分析的手法である安定同位体は、植物部位に関する情報を含みうる（Cernusak *et al.*, 2009）ので、両分析を同時に実施することで、互いに相補的に採食情報を知ることができるだろう。

NGSの技術革新は目まぐるしい。NGS（次世代シーケンシング）というが、現在第何世代にあたるのかフォローができないほど、新しい先進技術を搭載したNGS機器が次から次へ登場している（Heather & Chain, 2016; Shendure *et al.*, 2017）。たとえば、Oxford Nanopore Technologies社が開発したナノポアデバイスのシーケンサーは、非常に小型だ（Jain *et al.*, 2018）。製品のひとつであるMinIONは手のひらサイズで、コンピュータにUSB接続することで運用ができる。試料の複雑な前処理も不要で、DNAをインプットすることでそのまま塩基配列を決定できる。ポータブルで試料調整が簡便という特徴から、電力さえなんとかなれば、野外調査地でDNA解析をする時代もいよいよ近いと期待されている。スマートフォンに接続できるナノポアシーケンサーであるsmidgIONの開発も進んでおり、学校教育の現場などでの使用も想定されている。

DNAは生物の進化の原動力だ。38億年前に出現したあらゆる生物に共通する分子であり、細胞分裂時にはほぼ正確に複製される。その一方で、ごくまれに生じる変異の蓄積によって種分化が起き、現在の生物多様性を生み出した。こうした普遍的で頑健でありながら、小さな変異を許すという性質があるからこそ、分子生態学研究でもパワフルなツールとして利用できる。しかし目的がツールとしての利用であっても、DNAに向き合うときは常にDNAが生物進化の産物であることを意識したい。バーコード配列が分析できるのは、自然選択によって種を超えて保存されている領域と、中立進化によって変異に富む領域がゲノムに存在するからだ。DNAの解析技術は各段に進歩し、微量であっても検出できてしまう。それは同時に、コンタミネーションへの弱さも意味している。野外サンプルに対して素手で触ったり、マスクをせずにしゃべって唾液がかかったりすると、容易に研究者自身の細胞が混入する。ヒトと非ヒト霊長類は進化的に近縁がゆえに、DNA配列も非常に類似している。PCRでサルのDNAを増幅しようとしたら、進化的に一番近いヒトのDNAが増幅されてしまうということが簡単に起こりうる。野外でのDNA研究の注意点は、そのままDNAが本来持つ生物学的特性に結び付いている。

引用文献

- 1000 Genomes Project Consortium 2015: A global reference for human genetic variation. *Nature* 526: 68–74.
Abdelkrim J, Robertson B, Stanton JA, Gemmell N 2009: Fast, cost-

- effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* 46: 185–192.
- Adenyo C, Hayano A, Inoue E, Kayang BB, Inoue-Murayama M 2012: Development of microsatellite markers for grasscutter (*Thryonomys swinderianus*, RODENTIA) using next-generation sequencing technology. *Conservation Genetics Resources* 4: 1011–1014.
- Akomo-Okoue EF, Inoue E, Adenyo C, Hayano A, Inoue-Murayama M 2013: Development of microsatellite markers for blue duiker (*Cephalophus monticola*) using next-generation sequencing and cross-amplification in other duikers. *Conservation Genetics Resources* 5: 779–781.
- Amann RI, Ludwig W, Schleifer KH 1995: Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59: 143–169.
- Amato KR, Leigh SR, Kent A, Mackie RI, Yeoman CJ et al 2015: The gut microbiota appears to compensate for seasonal diet variation in the wild black howler monkey (*Alouatta pigra*). *Microb Ecol* 69: 434–443.
- Amato KR, Metcalf JL, Song SJ, Hale VL, Clayton J et al 2016: Using the gut microbiota as a novel tool for examining colobine primate GI health. *Global Ecology and Conservation* 7: 225–237.
- Andersen PS, Jespersgaard C, Vuust J, Christiansen M, Larsen LA 2003: Capillary electrophoresis-based single strand DNA conformation analysis in high-throughput mutation screening. *Hum Mutat* 21: 455–465.
- Arabidopsis Genome Initiative 2000: Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- Arandjelovic M, Vigilant L in press: Non-invasive genetic censusing and monitoring of primate populations. *Am J Primatol*.
- Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T et al 2011: Enterotypes of the human gut microbiome. *Nature* 473: 174–180.
- Blattner FR, Plunkett G, Bloch CA, Perna NT, Burland V et al 1997: The complete genome sequence of *Escherichia coli* K-12. *Science* 277: 1453–1462.
- Bohmann K, Evans A, Gilbert MT, Carvalho GR, Creer S et al 2014: Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol Evol* 29: 358–367.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD et al 2010: QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336.
- CBOL Plant Working Group 2009: A DNA barcode for land plants. *Proc Natl Acad Sci USA* 106: 12794–12797.
- Cernusak L, Tcherkez G, Keitel C, Cornwell WK, Santiago LS 2009: Why are non-photosynthetic tissues generally ¹³C enriched compared with leaves in C3 plants? Review and synthesis of current hypotheses *Funct Plant Biol* 36: 199–213.
- Chimpanzee Sequencing and Analysis Consortium 2005: Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437: 69–87.
- Choo JM, Leong LE, Rogers GB 2015: Sample storage conditions significantly influence faecal microbiome profiles. *Sci Rep* 5: 16350.
- Clayton JB, Vangay P, Huang H, Ward T, Hillmann BM et al 2016: Captivity humanizes the primate microbiome. *Proc Natl Acad Sci U S A* 113: 10376–10381.
- Constable JL, Ashley MV, Goodall J, Pusey AE 2001: Noninvasive paternity assignment in Gombe chimpanzees. *Mol Ecol* 10: 1279–1300.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL et al 2006: Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72: 5069–5072.
- Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L et al 2005: Diversity of the human intestinal microbial flora. *Science* 308: 1635–1638.
- Eggert LS, Maldonado JE, Fleischer RC 2005: Nucleic acid isolation from ecological samples--animal scat and other associated materials. *Methods Enzymol* 395: 73–87.
- Eisenberg D, Marcotte EM, Xenarios I, Yeates TO 2000: Protein function in the post-genomic era. *Nature* 405: 823–826.
- Flores R, Shi J, Yu G, Ma B, Ravel J et al 2015: Collection media and delayed freezing effects on microbial composition of human stool. *Microbiome* 3: 33.
- Folmer O, Black M, Hoeh W, Lutz R, Vriegenhoek R 1994: DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3: 294–299.
- Frantzen MA, Silk JB, Ferguson JW, Wayne RK, Kohn MH 1998: Empirical evaluation of preservation methods for faecal DNA. *Mol Ecol* 7: 1423–1428.
- Gorzelak MA, Gill SK, Tasnim N, Ahmadi-Vand Z, Jay M, Gibson DL 2015: Methods for Improving Human Gut Microbiome Data by Reducing Variability through Sample Processing and Storage of Stool. *PLoS One* 10: e0134802.
- Hale VL, Tan CL, Knight R, Amato KR 2015: Effect of preservation method on spider monkey (*Ateles geoffroyi*) fecal microbiota over 8 weeks. *J Microbiol Methods* 113: 16–26.
- Hashimoto C, Furuichi T, Takenaka O 1996: Matrilineal kin relationship and social behavior of wild bonobos (*Pan paniscus*): sequencing the D-loop region of mitochondrial DNA. *Primates* 37: 305–318.
- Hayakawa S 2008: Male-female mating tactics and paternity of wild Japanese Macaques (*Macaca fuscata yakui*). *Am J Primatol* 70: 986–989.
- Hayakawa S, Takenaka O 1999: Urine as another potential source for template DNA in polymerase chain reaction (PCR). *Am J Primatol* 48: 299–304.
- Hayakawa T 2015: Taste of chimpanzee foods. In Nakamura M, Hosaka K, Itoh N, Zamma K (eds): *Mahale Chimpanzees: 50 Years of Research*, Cambridge University Press, pp. 246–258.
- Hayakawa T, Sugawara T, Go Y, Udono T, Hirai H, Imai H 2012: Eco-Geographical Diversification of Bitter Taste Receptor Genes (*TAS2Rs*) among Subspecies of Chimpanzees (*Pan troglodytes*). *PLOS ONE* 7: e43277.
- Hayakawa T, Suzuki-Hashido N, Matsui A, Go Y 2014: Frequent expansions of the bitter taste receptor gene repertoire during evolution of mammals in the Euarchontoglires clade. *Molecular Biology and Evolution* 31: 2018–2031.
- Hayden EC 2014: Technology: The \$1,000 genome. *Nature* 507: 294–295.
- Heather JM, Chain B 2016: The sequence of sequencers: The history of sequencing DNA. *Genomics* 107: 1–8.
- Hernandez-Rodriguez J, Arandjelovic M, Lester J, de Filippo C, Weihmann A et al 2018: The impact of endogenous content, replicates and pooling on genome capture from faecal samples. *Mol Ecol Resour* 18: 319–333.
- Hofstetter JR, Zhang A, Mayeda AR, Guscar T, Nurnberger JI Jr, Lahiri DK. Genomic DNA from mice: a comparison of recovery methods and tissue sources 1997: *Biochem Mol Med* 62: 197–202.
- Human Microbiome Project Consortium 2012: A framework for human microbiome research. *Nature* 486: 215–221.
- Hyman ED 1988: A new method of sequencing DNA. *Anal Biochem*

- 174: 423–436.
- 井上英治 2015: 非侵襲的試料を用いた DNA 分析 —試料の保存, DNA 抽出, PCR 増幅及び血縁解析の方法について—. 霊長類研究 31: 3–18.
- Inoue E, Inoue-Murayama M, Vigilant L, Takenaka O, Nishida T 2008: Relatedness in wild chimpanzees: influence of paternity, male philopatry, and demographic factors. *Am J Phys Anthropol* 137:256–262.
- Inoue E, Takenaka O 2008: The effects of male tenure and female mate choice on paternity in free-ranging Japanese macaques. *American Journal of Primatology* 70: 62–68.
- International Human Genome Sequencing Consortium 2001: Initial sequencing and analysis of the human genome. *Nature* 409: 860–921.
- International Human Genome Sequencing Consortium 2004: Finishing the euchromatic sequence of the human genome. *Nature* 431: 931–945.
- Jain M, Koren S, Miga KH, Quick J, Rand AC et al in press: Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat Biotechnol*.
- Jayasankar, P., Anoop, B. and Rajagopalan, M 2008: PCR-based sex determination of cetaceans and dugong from the Indian seas. *Current Science* 94: 1513–1516.
- Kapusta A1 Suh A, Feschotte C 2017: Dynamics of genome size evolution in birds and mammals. *Proc Natl Acad Sci U S A* 114: E1460–E1469.
- Langergraber KE, Mitani JC, Vigilant L 2007: The limited impact of kinship on cooperation in wild chimpanzees. *Proc Natl Acad Sci U S A* 104:7786–7790.
- Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR et al 2008: Evolution of mammals and their gut microbes. *Science* 320: 1647–1651.
- Li H, Durbin R 2011: Inference of human population history from individual whole-genome sequences. *Nature* 75:493–496.
- Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ et al 2011: A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478: 476–482.
- Locke DP, Hillier LW, Warren WC, Worley KC, Nazareth LV et al 2011: Comparative and demographic analysis of orang-utan genomes. *Nature* 469: 529–533.
- Medini D, Serruto D, Parkhill J, Relman DA, Donati C et al 2008: Microbiology in the post-genomic era. *Nat Rev Microbiol* 6: 419–430.
- Moeller AH, Degnan PH, Pusey AE, Wilson ML, Hahn BH, Ochman H 2012: Chimpanzees and humans harbour compositionally similar gut enterotypes. *Nat Commun* 3: 1179.
- Moeller AH, Caro-Quintero A, Mjungu D, Georgiev AV, Lonsdorf EV et al 2016: Cospeciation of gut microbiota with hominids. *Science* 353: 380–382.
- Moeller AH, Foerster S, Wilson ML, Pusey AE, Hahn BH, Ochman H 2016: Social behavior shapes the chimpanzee pan-microbiome. *Sci Adv* 2: e1500997.
- Morin PA, Chambers KE, Boesch C, Vigilant L 2001: Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Mol Ecol* 10: 1835–1844.
- Morin PA, Moore JJ, Chakraborty R, Jin L, Goodall J, Woodruff DS 1994: Kin selection, social structure, gene flow, and the evolution of chimpanzees. *Science* 265:1193–1201.
- Mouse Genome Sequencing Consortium 2002: Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520–562.
- Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A et al 2011: Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332: 970–974.
- Mullis KB, Faloona FA 1987: Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* 155: 335–350.
- Nei M, Kumar S 2000: *Molecular Evolution and Phylogenetics*. Oxford University Press.
- Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW et al 2009: Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461: 272–276.
- Perofsky AC, Lewis RJ, Abondano LA, Di Fiore A, Meyers LA 2017: Hierarchical social networks shape gut microbial composition in wild Verreaux's sifaka. *Proc Biol Sci* 284: 20172274.
- Perry GH, Marioni JC, Melsted P, Gilad Y 2010: Genomic-scale capture and sequencing of endogenous DNA from feces. *Mol Ecol* 19: 5332–5344.
- Perry GH, Reeves D, Melsted P, Ratan A, Miller W et al 2012: A genome sequence resource for the aye-aye (*Daubentonia madagascariensis*), a nocturnal lemur from Madagascar. *Genome Biol Evol* 4:126–135.
- Pompanon F, Deagle BE, Symondson WO, Brown DS, Jarman SN, Taberlet P 2012: Who is eating what: diet assessment using next generation sequencing. *Mol Ecol* 21: 1931–1950.
- Prado-Martinez J, Sudmant PH, Kidd JM, Li H, Kelley JL et al 2013: Great ape genetic diversity and population history. *Nature* 499: 471–475.
- Prüfer K, Munch K, Hellmann I, Akagi K, Miller JR et al 2012: The bonobo genome compared with the chimpanzee and human genomes. *Nature* 486: 527–531.
- Queller DC, Strassmann JE, Hughes CR 1993: Microsatellites and kinship. *Trends Ecol Evol* 8: 285–288.
- Rhesus Macaque Genome Sequencing and Analysis Consortium 2007: Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 316: 222–234.
- Roach JC, Glusman G, Smit AF, Huff CD, Hubley R et al 2010: Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 328: 636–639.
- Rogers J, Gibbs RA 2014: Comparative primate genomics: emerging patterns of genome content and dynamics. *Nat Rev Genet* 15: 347–359.
- Ronaghi M, Uhlén M, Nyrén P 1998: A sequencing method based on real-time pyrophosphate. *Science* 281: 363–365.
- Sanders JG, Powell S, Kronauer DJ, Vasconcelos HL, Frederickson ME, Pierce NE 2014: Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. *Mol Ecol* 23: 1268–1283.
- Sanger F, Nicklen S, Coulson AR 1977: DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74: 5463–5467.
- Santana Q, Coetzee M, Steenkamp E, Mlonyeni O, Hammond G et al 2009: Microsatellite discovery by deep sequencing of enriched genomic libraries. *Biotechniques* 46: 217–223.
- Scally A, Dutheil JY, Hillier LW, Jordan GE, Goodhead I et al 2012: Insights into hominid evolution from the gorilla genome sequence. *Nature* 483: 169–175.
- Schoebel CN, Brodbeck S, Buehler D, Cornejo C, Gajurel J et al. 2013: Lessons learned from microsatellite development for nonmodel organisms using 454 pyrosequencing. *J Evol Biol* 26: 600–611.
- Sczyrba A, Hofmann P, Belmann P, Koslicki D, Janssen S et al 2017: Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software. *Nat Methods* 14: 1063–1071.
- Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J et al 2017: DNA sequencing at 40: past, present and future. *Nature*

- 550: 345–353.
- Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS et al 2017: Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 357: 802–806.
- Snyder-Mackler N, Majoros WH, Yuan ML, Shaver AO, Gordon JB et al 2016: Efficient Genome-Wide Sequencing and Low-Coverage Pedigree Analysis from Noninvasively Collected Samples. *Genetics* 203: 699–714.
- Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ et al 2016: Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems* 1: e00021–16.
- Srivathsan A, Sha JC, Vogler AP, Meier R 2015: Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Mol Ecol Resour* 15: 250–261.
- Sugiyama Y, Kawamoto S, Takenaka O, Kumazaki K, Miwa N 1993: Paternity discrimination and inter-group relationships of chimpanzees at Bossou. *Primates* 34: 545–552.
- Sun B, Wang X, Bernstein S, Huffman MA, Xia D et al 2016: Marked variation between winter and spring gut microbiota in free-ranging Tibetan Macaques (*Macaca thibetana*). *Sci Rep* 6: 26035.
- Swerdlow H, Gesteland R 1990: Capillary gel electrophoresis for rapid, high resolution DNA sequencing. *Nucleic Acids Res* 18: 1415–1419.
- Taberlet P, Waits LP, Luikart G 1999: Noninvasive genetic sampling: look before you leap. *Trends Ecol Evol* 14: 323–327.
- Tanabe AS, Toju H 2013: Two new computational methods for universal DNA barcoding: a benchmark using barcode sequences of bacteria, archaea, animals, fungi, and land plants. *PLoS One* 8: e76910.
- Tatsumoto S, Go Y, Fukuta K, Noguchi H, Hayakawa T et al 2017: Direct estimation of *de novo* mutation rates in a chimpanzee parent-offspring trio by ultra-deep whole genome sequencing. *Sci Rep* 7: 13561.
- Toju H, Yamamoto S, Tanabe AS, Hayakawa T, Ishii HS 2016: Network modules and hubs in plant-root fungal biomes. *J R Soc Interface* 13: 20151097.
- Tung J, Barreiro LB, Burns MB, Grenier JC, Lynch J et al 2015: Social networks predict gut microbiome composition in wild baboons. *eLIFE* 4: e05224.
- Vallender EJ 2011: Expanding whole exome resequencing into non-human primates. *Genome Biol* 12: R87.
- Vigilant L, Hofreiter M, Siedel H, Boesch C 2001: Paternity and relatedness in wild chimpanzee communities. *Proc Natl Acad Sci U S A* 98: 12890–12895.
- Wadman M 2008: James Watson's genome sequenced at high speed. *Nature* 452: 788.
- Watson JD, Crick FH 1953: A structure for deoxyribose nucleic acids. *Nature*. 171: 737–738.
- Weisburg WG, Barns SM, Pelletier DA, Lane DJ 1991: 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol* 173: 697–703.
- Wilson MR, Polanskey D, Butler J, DiZinno JA, Replogle J, et al 1995: Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. *Biotechniques* 18: 662–669.
- Zhang BW, Li M, Ma LC, Wei FW 2006: A widely applicable protocol for DNA isolation from fecal samples. *Biochem Genet* 44: 503–12.

Summary

Next-Generation Sequencing in Primate Molecular Ecology

Takashi HAYAKAWA^{1), 2)}*1) Department of Wildlife Science (Nagoya Railroad Co., Ltd.), Primate Research Institute, Kyoto University**2) Japan Monkey Centre*

Genetic investigation of wild primates are crucial to understand kinship, population diversity, phylogeographic patterns, and heritable factors of phenotypes. Traditional DNA technology using polymerase chain reaction (PCR) and Sanger sequencing have restricted the genome-wide analysis of primates, particularly due to the low quality and low quantity of noninvasive DNA samples obtained from wild individuals. Following the post-genome era, next-generation sequencing (NGS) technologies have provided a new paradigm in primate studies. NGS has enabled the genome-wide analysis of primate DNA using noninvasive samples, such as feces. Metabarcoding and metagenomics analyses using fecal samples provide information on food items and commensal microorganisms of the host animal. Here, I review a history of DNA sequencing technologies and examples of NGS studies in wild primates. Further, I discuss the effectiveness of NGS application to noninvasive samples.

Key words: DNA, noninvasive sample, next-generation sequencing (NGS), metagenomics

早川卓志

京都大学霊長類研究所ワイルドライフサイエンス（名古屋鉄道）寄附研究部門

公益財団法人日本モンキーセンター

〒484-8506 愛知県犬山市官林41-2 京都大学霊長類研究所

Takashi HAYAKAWA

Department of Wildlife Science (Nagoya Railroad Co., Ltd.), Primate Research Institute, Kyoto University

Japan Monkey Centre

Primate Research Institute, Kyoto University, Kanrin 41-2, Inuyama, Aichi 484-8506, Japan

e-mail: hayak.tak@gmail.com