

ゲノム解析技術の進展と課題 -巨大化する医学・生命科学分野の技術-

DEVELOPMENT AND CHALLENGES OF GENOME ANALYSIS TECHNOLOGY
-DNA SEQUENCING TECHNOLOGY BECAME LARGE-SCALE TECHNOLOGY-

荒内 貴子¹・井上 悠輔²・磯部 太一³・武藤 香織⁴

¹修士 (生命科学) 東京大学大学院 新領域創成科学研究科メディカルゲノム専攻 (E-mail: arachi@hgc.jp)

²博士 (社医) 東京大学医科学研究所 助教 公共政策研究分野 (E-mail: yinoue-kyt@umin.ac.jp)

³修士 (学際情報学) 東京大学医科学研究所 日本学術振興会特別研究員 公共政策研究分野
(E-mail: tisobe@ims.u-tokyo.ac.jp)

⁴博士 (保健学) 東京大学医科学研究所 教授 公共政策研究分野 (E-mail: krmt@ims.u-tokyo.ac.jp)

ヒトゲノム計画 (1990-2003) の進展によって DNA シークエンス技術は発展し、ヒトゲノム計画以後、数千～数万の塩基配列を産出する「次世代シーケンサー」が研究現場に導入されるようになった。生命科学や医学の領域では、シーケンシング技術の発展は、1 個の遺伝子から全ゲノム解析へと研究対象を変容させた。本研究の目的は、次世代シーケンシング技術の特徴を描写した上で、次世代シーケンサーを制御するために必要な研究環境を分析し明らかにすることである。さらに、ゲノム科学が抱える多くの問題について、国家レベルでの巨大プロジェクトに依存し、数多くの多様な領域の専門家の協働を必要とするような「巨大科学」であるという点から分析を講じた上で、それらの問題に対処するための方策を提示する。

キーワード：DNA シークエンス技術、次世代シーケンサー、ゲノム科学、ヒトゲノム計画、巨大科学

1. はじめに

生命科学や医学分野全体に大きな影響を与え、分子生物学の発展に寄与した発見や実験の技術は、これまでに数多く確立されてきた。たとえば、1953年に発見された DNA の二重らせん構造¹⁾、1977年に発表された塩基配列決定法^{2), 3)}、1983年に確立された PCR 法^{4), 5)}などが当てはまるだろう。これらの発見、技術の開発にはそれぞれノーベル賞が授与され、今日の分子生物学実験を支える概念や基盤技術となっている。

なかでも象徴的なのは、ヒトゲノム配列の全てを解読¹⁾することを目的として1990年から2003年に行われた「国際ヒトゲノム計画」の進行中に、塩基配列の解読を自動で行う装置であるシーケンサーの改良がなされたことである。塩基配列の解読は、生物の「設計図」といえるゲノムを読み解くことであり、シーケンサーは生命科学及び医学研究には欠かせない機器となっているが、近年、シーケンシング技術の発展が著しく、技術を開発し利活用するはずの人間が扱いきれないほど大きなデータをもたらす複雑で精密な機器が売り出された。

1960年代から、原子力産業をはじめ様々な科学技術が、莫大な資金と多くの人によって支えられる「巨大科

学」と化してきたこと^{6), 7)}、それによって人間が技術を制御できない状況が生じてきたことが指摘されており⁸⁾、また、近年では、第5世代コンピュータや大型放射光施設など巨大科学化してきた新たな科学技術分野が出現してきた状況も指摘されている¹⁰⁾。これまでの巨大科学の事例としては、宇宙開発におけるスペースシャトルや原子力など、物理的に大きく、目に見える形で専門家以外の一般市民にも分かりやすいものが主流であった。これらの科学技術は、日本の国力のアピールにつながるものが多く、国の威信にかけて行っているため研究予算配分などにおいては莫大な資金が投入されるものであり、巨大科学開発を進める過程においては人が機械のヒトコマとなって動く⁹⁾ことが指摘されている。以上のような、これまで巨大科学として捉えられてきた科学技術の「巨大科学としての特性」は、本研究が対象とするゲノム科学とも共通しており、国際ヒトゲノム計画も開始時に「医学生物学のビッグサイエンス」と呼ばれた¹¹⁾。また、国際ヒトゲノム計画は、大きな研究テーマの達成に対して多くの科学者が分業して関わる点から、新種の大型科学とも呼ばれることもある¹⁰⁾。

しかしながら、ゲノム科学には上記の原子力開発などの巨大科学と比較して特異的な側面がある。まず、ゲノ

ム科学の産物は、社会の側である一般市民からは一見して分かりづらいことが挙げられる。ゲノムデータが入った記録媒体を見ても、それが医療に直結するとは想像し難い。つまり、巨大な原子力施設やスペースシャトルの存在感と比較すると、今や、人1人分のデータがUSBスティック1本に納まるゲノムデータの存在感は社会の側からは地味に捉えられるのである。このように、目に見えづらい研究の営みと成果のために、一般市民にとってはゲノム科学が巨大化しているように理解されることは困難であろう。次に、ゲノム科学の成果は、どのように一般市民の生活に活かされるのかがわかりづらいという特徴も持つ。スペースシャトルも一般市民からすると距離は遠いが、スペースシャトルの飛行が成功すれば「宇宙へ行くことができる」という分かりやすさがある。しかし、ゲノム科学の成果が人類の福祉向上につながる可能性やその過程は、専門的な解釈を要し、理解することが難しいといえる。

以上のように、ゲノム科学は、先行する他の巨大科学と共通点もあるが、成果や研究の捉え方をめぐる多くの相違点がある。このような背景から、その後のシーケンス技術の発展やそれを利用した大規模なゲノム科学は、いわゆる「巨大科学」という視点では注目されず、技術を利用するゲノム研究者のなかでシーケンス技術の応用の仕方や機器の取り扱い方が論じられるに留まっている。新型のシーケンス技術の登場によってゲノム科学が進む一方で、研究者が機器を扱いきれていない状況が生じている現在、新たなシーケンス技術を可能にしている機器の取り扱い方を再考し、改めて「巨大科学」という視点でシーケンス技術を捉えなおしてみたい。

2. ねらい

新型の機器である、いわゆる「次世代シーケンス技術 (next-generation sequencer)」¹¹⁾は、原理・解析ともに従来のシーケンス技術とは全く異なるが、どのような経緯で従来のシーケンス技術とは異なる機器が登場したのだろうか。そして、どのように異なり、どのような影響を研究現場に及ぼしているのか。また、1960年代頃から、莫大な資金と多くの人間を動員するような科学は「巨大科学」と呼ばれはじめ^{6,7,8,9,10)}、主として原子力産業を指したが、シーケンス技術の著しい発展とそれを利用した大規模なゲノム科学を考えると、ゲノム科学も、巨大科学の仲間入りを果たしたと考えることができる。

そこで、本論文で対象とする次世代シーケンス技術を利用した、2005年以降に発売された、電気泳動をおこなわずにDNA断片の検出を大量並列に解析し、塩基配列の決定をおこなう解析機器と定義する。そのうえで、本論文では、ゲ

ノム解析技術の中心といえるシーケンス技術の発展過程を明らかにするとともに、次世代シーケンス技術を利用したゲノム解析技術がどのような性質をもつ技術で、利用するにあたって何が必要になるのかを明らかにすることを目指した。また、高速かつ大量にゲノム配列を解読する次世代シーケンス技術の利用によるゲノム解析がどのような課題を抱えているのかを、巨大科学化という視点を交えて考察したい。

本研究の分析視点としては、研究労働環境や研究のコスト最適化など、シーケンス技術の発展過程とそこでの研究者と機械との関係性に焦点を置いた。そのため、「個人のゲノム解析の是非」をはじめとする倫理的・法的・社会的諸課題については分析の対象外とする。

研究手法としては、シーケンス技術の発展過程を明らかにし、従来のシーケンス技術と次世代シーケンス技術を利用したシーケンス技術の違いを検証するため、ゲノム研究者によるシーケンス技術に関する論説や論文をもとに文献調査を行った。このほかに文献調査の補足として、文献調査から得られた情報の解釈や、研究現場における次世代シーケンス技術に関する諸課題の確認のため、次世代シーケンス技術を保有する研究施設のゲノム研究者2名へのインタビューも行った。

3. シーケンス技術の発展

3.1. これまでのシーケンス技術

そもそも塩基配列決定法は、いつ誰が開発し、どのように進展してきたのだろうか。開発された当時の主な塩基配列決定法は2つあり、1977年にサンガーとギルバートによって、それぞれ異なる原理の手法が確立された。発表当初はギルバートが確立した手法の方が利用されていたが、次第にサンガーが確立した手法(サンガー法)がよく利用されるようになった¹²⁾。サンガーは、1960年代から塩基配列決定法の開発に取り組んでおり、その後利用されなかったが、73年と75年に塩基配列決定の手法に関する論文を発表している^{13),14)}。75年の論文には既に、「a simple and rapid method for determining nucleotide sequences」という表現をしており、この時点で「簡単に速い塩基配列決定」を目指していたことがわかる¹⁴⁾。後に、サンガーはこのころを振り返り、*Nature*誌に寄稿している。塩基配列決定法の開発を行い始めた理由として、「DNAについて多くのことが発見され、我々は生物にとって必要な情報はすべて塩基配列に含まれているのだということに気がついた」と述べており、塩基配列の決定¹⁵⁾は、生物の理解と医学への貢献を目指していたとしている¹⁵⁾。

現在も利用されているサンガー法だが、当時は、全て

の煩雑な実験を手作業で行い、数か月かけてようやく1000塩基決定できる程度の実験であった¹⁶⁾。当時の実験書を参照すると、Table 1に示したように、塩基配列を知りたいサンプルの前処理から配列の検出まで少なくとも3, 4日かかることがわかる。そこで、1980年には、和田昭允が煩雑な塩基配列決定法の工程の自動化を提唱した。この技術の開発は、科学技術庁(当時)の科学技術振興調整費によって運営される国家プロジェクト「DNAの抽出・解析・合成技術の開発に関する研究(和田プロジェクト)」として採用され、塩基配列決定を自動化する試みが行われた¹²⁾。

和田が投稿した塩基配列決定の自動化に関する論文は、「Analytical Biochemistry」という雑誌では「DNA自動解析装置は専門家による解析のレベルには達していない」という理由で返却されたが、「The Review of Scientific Instruments」という雑誌には「DNAの塩基配列に必要な退屈な工程を自動化し、標準化する機器を報告している。疑いもなくこれからの分子生物学、遺伝学の研究に有用な役割を果たすものである」と判断され、無修正で掲載となった¹⁷⁾。この背景には、当時の分子生物学実験には熟練した技術が必要であり、和田は日本国内で「機械が解読の名人(熟練者)にかなうはずがない」「人間の手で

やれるのに、なんで機械にやらせるのだ」という批判を浴びた¹⁷⁾。

こうしたなか、1986年には米国で、世界で初めてのシークエンサーが発売された。Table 1に示したように、撮影と検出の部分を自動化したものだ。このころ、日米の貿易摩擦の対策として発された、中曽根康弘首相(当時)の「バイ・アメリカン(米国製品を買おう)」という掛け声に押され、シークエンサーは補正予算がついて導入された。しかし、シークエンサーは「研究室の片隅でホコリを被るケースが多」く、それは品質の悪さではなく、使い勝手の悪さによるものだった¹⁶⁾。

1987年には和田プロジェクトの一員であった日立製作所の神原秀記が自動でシークエンスが行える機器を開発した¹⁶⁾。その後、1993年には、平板ゲルを利用した方法からキャピラリーという細いガラス管を利用した機器が開発され、サンガー法を原理としたシークエンス技術は、年々高速化・ハイスループット化し¹⁸⁾、一度に得られる配列データを増やしてきた¹⁸⁾。前処理以外はすべて自動化されたことをTable 1に示した。はじめは、ひとつのサンプルの実験を自動で行うことが目的だったが、それが達成されると、キャピラリーを増やして同時に違うサンプルの実験も行うようなシステムが開発された。

	前処理 (テンプレートDNA の増幅・調整)	配列決定のための 化学反応 (ジデオキシ反応)	DNA断片を長さごと に分離 (電気泳動)	検出		一度の実験または 稼働で解読できる 塩基数
				撮影	検出	
塩基配列決定法† (1977-)	手動 (6時間程度)	手動 (2時間程度)	手動 (半日以上)	撮影 手動 (1日)	検出 手動 (作業時間は 作業による)	200塩基程度
スラブ式 シークエンサー†† (1986-)	手動 (6時間程度)	手動 (2時間程度)	電気泳動を始める ところまでは手動 (4時間程度)	自動 (1日)		600-800塩基
キャピラリー式 シークエンサー††† (1993-)	PCR装置を利用するが、 それ以外の実験は手動** (6時間程度)			自動 (2時間程度)		600-800塩基 最大384サンプル

	前処理 (試料DNA 調整・増幅)			配列決定のための 化学反応・撮影・検出		一度の稼働で解読 できる塩基数
次世代 シークエンサー†††† (2005-)	手動** (6時間程度)			自動 (3日-)***		10億塩基-***

†サンガーが発表した塩基配列決定法の論文と遺伝子工学の実験書を参考にした^{21), 19), 20)}。††遺伝子工学の実験書を参考にした^{21), 22)}。†††3社の次世代シークエンサーのうちのひとつを例としてあげており、次世代シークエンサーを利用した実験法を紹介している実験書等を参考にした^{23), 24)}。*キャピラリー式シークエンサーの時代には、PCR (polymerase chain reaction) という、生体外での遺伝子増幅技術がマリスによって確立され、シークエンス技術にも利用され始めた。**最近では、前処理を自動化も行われ始めている。***表中に記載したのは、最初の次世代シークエンサーの解読できる塩基数と稼働時間であり、改良され続けている。

Table 1 シークエンス技術の発展に伴う作業工程の変化と解読できる塩基数
上の表はサンガー法の塩基配列決定法が自動化される前後の作業工程を示しており、下の表は次世代シークエンサーを利用した場合の作業工程を示している。作業時間については、実験書に記載されている作業時間や待ち時間などを参考にしたほか、実験を行っている研究者から聞いた実際の作業時間を参考にした。

3.2. 国際ヒトゲノム計画とシーケンス技術の進展

従来、生命科学分野や医学分野では、1つの遺伝子に関する研究がなされてきたが、1980年代前半から、多くの研究者が1つ1つの遺伝子を個別に調べているのは、生命現象を理解する上で限界があることに気づき始めた²⁵⁾。1985年、米エネルギー省がサンタフェのロスアラモス国立研究所でヒトゲノム配列の解読に関するワークショップを開いている。この会合で、生物物理学者のチャールズ・デリシがヒトゲノム計画の提案をし、翌86年にはエネルギー省から5.3百億ドルをかけたパイロット研究が発表された²⁶⁾。また、国立衛生研究所(National Institute of Health)も、1986年にシンポジウム「ホモサピエンスの分子生物学」をコールドスプリングハーバー研究所にて開催した^{27),28)}。DNA二重らせん構造を発見したジェームズ・ワトソンは、同研究所の所長として、このシンポジウムの中で「ヒトゲノム計画」の特別セッションを設け、米国内の研究者を主導する役割を果たした²⁸⁾。ワトソンを支援した著名ながん研究者であるダルベッコは、「がんについてより詳しく知るためには、ゲノム全てを解読するべきだ」とサイエンス誌に寄稿している²⁹⁾。

このシンポジウムを契機として、ワトソンは世界の先進国に対しても強く「協力」を呼びかけていく³⁰⁾。その結果、1988年には、国際協力の連携組織としてのHUGO (Human Genome Organization) の設立準備の会合がスイスで開かれた。HUGOは、研究者が自主的に国際的連携・推進のために設立した組織であり、国際ヒトゲノム計画にかかる研究者たちの強い意気込みの表れだと、計画に関わってきたゲノム研究者の一人である榎は述べている²⁷⁾。

こうして、ヒトゲノム配列解読への気運が高まり、1990年に「国際ヒトゲノム計画」がスタートしたが、その進行に伴い、サンガー法を原理としたシーケンサーは、サイズの大きいヒトゲノムを解読するために、より速く大量に配列を解読することを目指して、複数のサンプルのシーケンスができるように改良が進んだ。ヒトゲノムは、長いDNA鎖を端から解読していくのではなく、短い配列の断片に切断し、そのひとつひとつをシーケンスし、後からつなぎ合わせる作業(アセンブル)をしていたので、シーケンサーがなるべく多くのサンプルを処理できることは重要だった。また、一度解読した配列データは、粗配列データと呼ばれ、不確定の塩基を含んでいるために、最終配列にはならない¹⁸⁾。従って、同じ配列を何度も解読し、不確定の塩基データを出してしまう読み取りエラーの可能性をなくす必要があるため、シーケンサーの高速化はヒトゲノム配列解読には欠かせない要素のひとつだった。つまり、国際ヒトゲノム計画が、シーケンサーの技術革新に大きく貢献した契機

となったといえるだろう。

1980年代は、ひとつの遺伝子さえ見つけることが難しい時代であり、およそ30億塩基対からなる膨大なヒトゲノム配列を読み解くことは100年かかっても不可能と考えられてきた²⁵⁾。しかし、多くの資金とシーケンサーの投入やコンピュータの発達により、計画の終結は早まった。サンガー法に代わる方法の開発が進まなかったため、上記のように、ヒトゲノム配列解読の速さはサンガー法のシーケンサーがどれだけ多くのサンプルを解読できるかにかかっていたが、計画の進行中にキャピラリー式のシーケンサーが登場し、著しく改良された。国際ヒトゲノム計画では、そうしたシーケンサーを数十から数百台を並列して解読するという大規模センター方式による解読が世界共通となった¹⁸⁾。15年間の予定でスタートした計画は、10年目の2000年にドラフト配列を発表し、2003年には完全解読を達成したため、2年短縮して終結した。

2003年以降は研究者がヒトゲノム配列を研究に用いることが可能になったため、その配列を参照配列として、ゲノム上でどこがどのように変異しているかという発見あるいは確認や、新たな遺伝子の発見などこれまで難しいと思われた解析が可能となった。しかし、米国が中心となり、6カ国(米、英、仏、独、日、中)、24機関が協力し、潤沢な資金と数百台のシーケンサーが投入されている状態でも^{16),17)}、サンガー法では13年かかったのであるから、ヒトゲノム同士を丸ごと比較するような解析(Whole-genome解析)をサンガー法で個々の研究者が行うのは不可能に近いといえる。現在では、一度の稼働で最大384個のサンプルを扱うことができる機種がでて^{18),25)}いるが、384サンプル一度に解析できるシーケンサーを1台利用してヒトゲノム配列の解読を行ったとすると、サンガー法では、1サンプルにつき800塩基ほど解読できるので、単純計算で約30年要する。配列の解読は機器が自動的に行ってくれるが、機器にかけるまでのサンプルの調整や、読み取りエラーの発生に対処し、その後、ばらばらに切断して解読した配列データをつなぎ合わせる(アセンブル)ために繰り返し配列を解読することを考慮すると、より時間がかかる。

4. シーケンス技術の現状

4.1. 次世代シーケンサーの特徴

網羅的な解析が行えるように、国際ヒトゲノム計画後も高速かつ大量に配列データが得られるようなシーケンス技術の原理の開発が進められた。ヒトゲノム計画においては、サンガー法の効率を上げるだけではヒトゲノムを解読することは難しいと考えられ、シーケンスの

スピードを画期的に挙げる新しい方法が必要ではないかと考えられていた²⁴⁾。また、2004年には米国ヒトゲノム研究所 (National Human Genome Research Institute) がヒトの Whole-genome 解析を低コストかつ効率的に行えるシーケンシング技術の開発を目指した 3800 万ドルの助成金を発表した³⁰⁾。そうした背景から開発された機器が、2005 年以降に発売された「次世代シーケンサー」である。サンガー法を原理としたシーケンサーでは、キャピラリーによる電気泳動(断片の長さ別に分離する方法)の工程が含まれていたのだが、次世代シーケンサーでは、電気泳動をなくし、その時間を短縮するとともに、1 回に解析できるサンプル数を飛躍的に高めた²⁵⁾。

それでは、3.2 で紹介した従来のシーケンサーと次世代シーケンサーはどのように異なるのか。2005 年から 2007 年にかけて 3 社から発売された次世代シーケンサーは³³⁾、「1G (ギガ) シーケンサー」と呼ばれ、10 億塩基の配列を解読できるものであった。サンガー法を原理とした従来のシーケンサーは並列して配列を解読できるサンプルは長らく 384 個が限界であり続けたのに対し、次世代シーケンサーでは約 4000 万個の配列を同時に解読でき、その限界を突破することができた。ヒトゲノム配列が約 30 億塩基なので、一度の稼働でその 3 割を解読できることになる。つまり、およそ 10 日で一人分のヒトゲノム配列が得られる計算になり、サンガー法とは比にならない速さである。したがって、次世代シーケンサーは、解読速度という観点では、ヒトゲノムの全領域を参照配列と比較することも可能にしたといえ、現在もさらなる改良でハイスループット化が進んでいる。

また、コストの面でも、従来のシーケンサーと比較すると塩基配列の解読にかかるコストは低くなっている。1 塩基を解読するために必要なコストは、従来のシーケンサーよりも、次世代シーケンサーによる配列の解読のほうが安価だ²⁴⁾。

他方、次世代シーケンサーが、従来のシーケンサーより劣る点として、まず精度の問題がある。ヒトゲノム中には、個人間の違いである一塩基多型 (SNP: single nucleotide polymorphism) が 0.1%程度存在しているが³²⁾、読み取りの精度が低いと、参照配列と異なる塩基が SNP なのかエラーなのか見分けがつかない。そこで、同じサンプルを何度も配列を解読し比較することで、SNP とエラーを区別する必要がある。また、サンガー法で得られる配列データは長いので、配列データ同士の重なる部分も多く、ばらばらに断片化したゲノム配列をつなぎ合わせやすかったが、次世代シーケンサーでは、リード長(一度に読まれる DNA 断片の長さ)が短く、配列データの重なる部分を手掛かりにつなぎ合わせる作業が困難となった。2007 年以降に発売された次世代シーケンサーでは、リード長を延伸するよう改良されているものの、

エラー率の引き下げには繋がっていない^{34),35)}。

一度の稼働にかかる時間も、サンガー法のシーケンサーに比べて長い。従来のシーケンサーでは、24 時間以内に反応が終わり、データが得られたが¹⁸⁾、次世代シーケンサーでは数日単位の反応となる。そのため、サンプル調整からデータが得られるまでのスパンが長く、研究者の都合に合わせて実験するというよりは、機器の稼働状況に合わせて実験することになる。次世代シーケンサーには、サンプル調整と操作のみを行う人材を配置しており、機器を稼働させ、その稼働状況を適宜確認している研究室もあるようだ。

また、次世代シーケンサーで得られたデータの解析は、従来よりも難しいという特徴がある。サンガー法のシーケンサーから得られるデータは手元のパソコンで扱える程度のデータ量であり、特別な情報解析のための設備は必要ないが、次世代シーケンサーから得られるデータは、保存・解析ともに、手元のパソコンでは不可能であり、スーパーコンピュータなどの大型の計算機の利用が必要になる^{24),36)}。そのためには、設備だけではなく、高度な情報解析のための知識が必要となる。習得を要する情報解析の技術は、大別して 2 つある。まず、計算機そのものを扱えることが必要になり、Linux 等の OS を利用しなければならないため、Windows 等の視覚的にわかりやすい OS とは異なり、コマンドラインで全ての処理をしなければならない。さらに、配列の解析は、エクセルなどの汎用ソフトウェアを利用できないため、Perl 言語などのプログラミング言語を駆使して集計・整理することになる。これまで分子生物学的手法を用いて実験主体の研究を行ってきた研究者には、あまりなじみのない分野であり、新たに上記のような情報解析の手法を学ばなければならない。

4.2. 日本における次世代ゲノム解析技術

それでは、日本ではいったいどのくらいの次世代シーケンサーが稼働しているのだろうか。日本の次世代シーケンサーの導入状況は、21 台であると発表されているが、実際には、2009 年の補正予算により、より多くの研究機関に導入されていると考えられている。公表を控えている研究機関が多いため、導入状況に関してはわからない部分が多いが、その導入形態は、米国や中国、英国のように大きな研究機関に集中的に導入されているというのではなく、研究施設あるいは研究室ごとに分散的に導入されており、各々で機器の管理・運用をしているという形態である。しかし、解析には大型の計算機を必要とし、たとえばスーパーコンピュータは、限られた場所にしかなく、東京大学や東京工業大学、理化学研究所などに設置されている。マンハイム大学、テネシー大学、ローレンス・バークレイ米国国立研究所の研究者らに

よるスーパーコンピュータの速度による格付けを行っているデータベースである TOP500 (<http://www.top500.org/>) を参考にすると、2013年6月現在、100位以内のランキングに入っている日本のスーパーコンピュータは10台である。この中には大学用や研究用のためのものだけでなく産業用のためのスーパーコンピュータも含まれるため、全てが研究用ではない。さらに、スーパーコンピュータはゲノム科学だけに利用されているわけではなく、他の多くの自然科学分野で利用されている。この点を考慮すると、次世代シーケンサーの正確な台数はわからないが、利用可能なスーパーコンピュータの台数を上回っていることは推測できる。こうした状況下では、次世代シーケンサーを稼働しやすいとは言えない。

日本における分散的な機器の導入は、中国のように集中的に多くの次世代シーケンサーが導入されている国に比べると、何度も稼働を繰り返すような Whole-genome 解析は向かないと言えるが、その他の応用的な解析を行っている研究者もいる。次世代シーケンサーを導入している研究施設あるいは研究室は2009年以前にもあり、うまく稼働できているところもあるので、日本からも次世代シーケンサーを用いた研究の成果が論文として発表されていないわけではない。しかし、2010年に行ったゲノム研究者へのインタビューによると、2009年以降に機器の導入を決めた研究施設・研究室では、導入後すぐにはうまく稼働できず、次世代シーケンサーを扱いきれていないところも多い。

5. 考察

5.1. 従来のシーケンサーと次世代シーケンサーの比較

4.1で分析したとおり、次世代シーケンサーには、大量かつ高速に配列データが得られるという以外にも手動で行っていた塩基配列決定法や従来のシーケンサーとは異なる特徴がある。それは、エラーの多さと配列データの長さであり、次世代シーケンサーが行っていることは、塩基配列「決定」法と呼べるのかという疑問が生じる。サンガーやギルバートが確立した技術は“a new method for determining nucleotide sequences”と表現されており、「塩基配列決定法」と訳されている。この実験でも、実験者による「読み取りエラー」と呼ばれるエラーが存在し、それを減らすために相補鎖の塩基配列も合わせて読み取り、それらが合致しているかを確認しながら塩基配列を「決定」していったようだ¹⁹⁾。つまり、「読み取り（解読）」と「決定」が極めて近い作業だったと考えられる。しかし、次世代シーケンサーでゲノム配列を「決定」しようとする、繰り返す同じサンプルの配列を解

読し、それらを比較してエラーをなくしたり、得られた短い配列同士をつなぎ合わせたりするので、知りたいゲノム配列の約50倍の量のデータが必要になる。現在の技術では、エラーを減少させることや長い配列を扱えることも目指した改善がなされているが、基本的には、得られるデータ量を増やし、より多くより速く解読することで「決定」しようとしている。また、次世代シーケンサーで解読したものを確実な配列とするため、従来のサンガー法のシーケンサーで確認するという声も聞く。次世代シーケンサーが行っていることは塩基配列を「解読」するところまでであり、「決定」には4.2で触れたような設備や人材や従来のシーケンサーの助けが必要になる。これらの現象から、次世代シーケンサーのシーケンス技術においては、サンガー法に比べ「解読」と「決定」の乖離が起き、「解読」量が肥大化したと言える。

次世代シーケンサーと従来のシーケンサーには、共通点もある。これまで分析してきたとおり、シーケンス技術は、和田によって自動化が提唱され、現在でも利用されているサンガー法のシーケンサーが開発された。和田は塩基配列決定法の実験を人間の手ではなく機械にやらせるべきではないとの批判を浴びるが、塩基配列決定法の自動化は進み、国際ヒトゲノム計画という国際的で大規模なプロジェクトを遂行するためにその技術は著しく進展した。しかし、日本国内においてアメリカから大量に導入されたシーケンサーは、使い勝手の悪さと実験は人の手でやるものだという考え方から、しばらく利用されない状態が続いた。この状況は4.2で述べたとおり、現在の次世代シーケンサーが置かれている状況と共通する。

次世代シーケンサーの場合には、機器が行う工程は人間にはできない高度なものなので、実験は機械にできるものではないという実験科学者の職人的な考えではなく、3.3で分析したようなコストや設備、人材の問題を解決しなければ利用できないという使い勝手の悪さが稼働させられない理由になっていると考えられる。ここで、20年前も現在も、国の予算で導入した機器が使い勝手が悪いために稼働までに時間がかかったという共通点は、明記しておくべきだろう。

「使い勝手の悪さ」はどこからくるのか。新しい機器を利用するには、当然その機器を稼働させるための知識と技術が必要になり、利用者がそれらを十分に習得していなければ、機器は利用されにくくなる。シーケンサーを日本に導入する際にも、そういった基盤が整っていなかったのではないだろうか。20年前も現在も、米国製の機器を急激に導入した経緯があり、稼働に必要な知識や技術の整備が追いついていないということは、20年前に経験したことが現在に活かされていないことを示して

いるのではないだろうか。研究活動が円滑に行えるように導入後に機器が放置されることは避け、稼働できる環境を早急に整える必要がある。

シーケンス技術は今後もますます進展し、価格低下とダウンサイジングも同時に進行し、導入を検討する研究機関が増えることが予想されるが、導入に先んじてその稼働環境を整えることが望ましいことには変わりはない。次世代シーケンサーを保有する研究機関でユーザーミーティングなどを発足させて、次世代シーケンサーの開発・販売企業とも連携しながら、導入前に稼働環境を維持するための方策を情報共有すべきである。

5.2. 次世代シーケンサーが及ぼす研究現場への影響

次世代シーケンサーを利用したゲノム解析が従来の分子生物学や分子遺伝学と異なる点として、機器自体のコストとランニングコストが従来よりも高い点、稼働には専門知識（分子生物学・情報解析）を持った人材と高性能な大型計算機が必要である点があげられる。ここでは、それらの違いが、次世代シーケンサーの研究機関への導入という観点から研究活動に与える影響を考察する。

まず、機器自体のコストであるが、表で示した通り、3社の次世代シーケンサーは全て高価であるうえ、次世代シーケンサー自体の他に、データ保管・解析用の計算機が必要である点や、新機種やバージョンアップが短期間のうちに発表されるので、対応しなければならない点など費用がかさむ要因がいくつもある。

さらに、機器を稼働させていくうえで考慮すべきなのは、1塩基あたりのコストよりも1回の稼働あたりのコストである。1塩基あたりのコストが安価でも、一度稼働させると結局はサンガー法のシーケンサーを稼働させるよりもコストがかかり、潤沢な研究費を有している研究室でなければ、この機器を稼働させることは難しい。

この課題を克服するための解決策として、フローサイトメーターや電子顕微鏡などのように、次世代シーケンサーを各研究機関が共同利用できる機器に位置付けることが考えられる。そうすれば、一度の稼働で様々なサンプルと一緒に解読できるので、コストは利用した分だけ払えばよいことになり、研究室にかかる負担が少なくなるだろう。次世代シーケンサーを購入する予算がない研究室でも、共同利用が可能であれば、次世代シーケンス技術にアクセスしやすくなる。コスト面以外の問題として、次世代シーケンサーはすぐに稼働できる機器ではなく、様々な設備や人材が必要である。まず、誰が、機器の操作をするのかという問題が生じる。先に述べたように、次世代シーケンサーは一度の稼働で大量の塩基配列を解読でき、異なるサンプルを同時に解析す

ることも可能だ。共同利用の場合には、全て外部からの持ち込みサンプルの稼働になるので、サンプルに関わらず機器を操作するための人材が必要だろう。実際には、若手研究者を次世代シーケンサーの稼働要員として雇う事例が目立っており、日々他人のサンプル調整のサポートをし、次世代シーケンサーを稼働させることに追われ、自分の研究に必要な時間を確保しにくくなっているとの声も挙がっている。しかし、そうした人材は、分子生物学の実験ができ、次世代シーケンサーに精通していなければならないが、サンプル調整の技術さえあれば研究者である必要はない。実験補助業務としての人材育成が急務である。

また、スーパーコンピュータのような高価で大型の計算機は、研究室単位で簡単に用意できるものではなく、全ての研究機関で導入しているわけでもない。施設内外の計算機を利用できる基盤が整ったとしても、解析を行う人材がいなければ、次世代シーケンサーは稼働できない。

近年、バイオインフォマティクスという分野が確立されたが、この分野はデータ解析や新しいアルゴリズムの開発を行う分野であり、オミックス研究が盛んに行われるようになってからますます必要な分野となってきた。ゲノム情報解析に必要な計算機の台数は指数関数的に増え、それを動かせる人材を確保するために、バイオインフォマティクス教育体制も再検討が必要との指摘がある³⁶⁾。

次世代シーケンサーから得られたデータの一次解析には、日々必ず行われる作業として、配列情報がゲノムのどの位置のものかを調べるための参照ゲノムへのマッピングや長い配列を決めたいときには配列情報同士をつなぎ合わせるアセンブルなどの作業が発生する。この作業を行う人材として、次世代シーケンサーを操作する若手研究者と同様に、バイオインフォマティクスを学んだ若手研究者に期待が集まったり、実際に投入されたりする傾向にある。しかし、日々の作業に忙殺されて研究できない状況を作り出すのは、研究者の雇用のあり方として疑問である。やはり、計算機の共同利用や受託解析システムの確立が必要であり、研究者以外の人材育成が急務であろう。

以上のように、次世代シーケンサーを利用した研究を行う際には、どの部分が日々行われる作業で、どの部分が研究活動にあたるのかを検討しなおす必要に迫られているといえるだろう。作業は、実験補助業務を担う人材に任せ、研究者は研究活動に時間を割くべきである。具体的には、サンプルの調整までは研究者が行い、次世代シーケンサーの稼働は、実験補助者に任せることが可能だ。また、データ解析における日々の反復作業をシステムエンジニア職の人材に任せることで、研究活動と

しての情報解析やデータの解釈に時間を割くことができるようになる。

日本の医学研究の総合戦略の策定を担う組織として2015年から本格始動する日本医療研究開発機構構想では、大規模全ゲノムシーケンス解析基盤として、「セントラルゲノムセンター」を配置することが決定されている³⁷⁾。また、全ゲノムシーケンス解析を受託する企業も増えている。今後、設備と人材の整った解析拠点が絞り込まれ、研究機関においてもシーケンス解析の外注が増加することが見込まれる。こうした展開により、本論文で指摘した課題が解決されるかどうかは、注視する必要がある。

なお、次世代シーケンサーを操作する場合も、得られたデータをスーパーコンピュータで解析する場合も、ヒトゲノム配列情報というデータをどう取り扱うべきなのかは課題である。臨床検体から得たヒトゲノム配列情報はその人個人の「設計図」と表現されることが多いが、実際の配列データは高度な技術と設備がなければ解読できないうえ、別途、参照情報のデータベースがなければ、個人の特定には至らないという性質を持つ。今後、個人遺伝情報が氾濫し、参照情報データベースとの突合が容易になる時代に備え、2013年に改正された文部科学省・厚生労働省・経済産業省「ヒトゲノム・遺伝子解析研究に関する倫理指針」では、個人情報除かれたヒトゲノム配列情報の取り扱いについても、安全管理措置を要求している。ただし、指針には具体的な水準は明記されず、運用が研究者に任されている点には問題がある。多くの研究機関でそれぞれの安全管理措置の内容を共有しながら、一定の標準化を進めるべきである。

5.3. 巨大科学化したゲノム科学

シーケンス技術の変化は、ゲノム研究をどのように変容させたのか。国際ヒトゲノム計画や、従来とは異なる原理で成り立つ次世代シーケンサーの登場は、単に科学的・技術的な進展としてではなく、ゲノム解析技術の巨大科学化の過程と捉えることが可能だろう。

巨大科学に関する文献は、1960年代から1970年代にかけて重点的に発表されている。巨大科学は、富（巨大な科学研究や技術開発を行うために必要な資金）の集中と知（多くの広範な分野の科学者）の集中と力（高度な情報管理と中央集権的な研究開発システム）の集中が必要となる科学であり、そこで扱われる技術や物質がひとりの人間にとっては、計り知れない破壊的な力を持つという特徴を持つ³⁸⁾。巨大科学と考えられる最初の事例は1945年に米国で行われたマンハッタン計画と考えられており³⁹⁾、具体的には、原子力開発利用、宇宙開発⁴⁰⁾などを指すことが多い。

本論文で明らかにしたように、国際ヒトゲノム計画に

は多くの国がかかわり、多大な資金とシーケンサーが投入されたことから、巨大科学と言えるだろう。また、次世代シーケンサーの登場以来、ゲノム解析技術における情報解析の高度化が進み、ますますゲノム解析技術がゲノム研究者だけのものではなくてきていることがわかる。これも、巨大化が進み、個々の研究者が担う役割が細分化された結果ではないだろうか。高木は、こうした状況を「機械のひとコマになった個人」と表現し、役割が細分化されると、研究の全体性やその意味について、十分に見通すことができなくなり、マシンに振り回されることによって人間性をすり減らすと述べている⁸⁾。つまり、次世代シーケンサーが研究活動に導入されたことで、ゲノム科学の巨大化はさらに加速され、常態化した状況を生み出したのである。

より多くの配列を短時間で得れば、研究のスピードも上がるので、シーケンス技術の向上が望まれており、2005年に次世代シーケンサーが登場した。しかし、次世代シーケンサーは高額であり、サンガーシーケンスよりも多くの作業（実験→シーケンサーに読ませる→データの解析・整理）や次世代シーケンサーのための環境整備（機器自体、スパコン、人材）が必要となり、研究室で簡単には使いこなせない事態を招いた。

これまでも述べてきたように、次世代シーケンサーを中心としたヒトゲノム計画は巨大科学として位置づけることが出来、その巨大化の特徴と過程は大きく分けて以下の3点に集約できる。

まず1点目として、次世代シーケンサー導入の環境整備と機器の購入には多額の資金が必要になることがあげられる。第2点目は、シーケンス作業で出てくるデータがサンガーシーケンスよりも遥かに膨大であり、その膨大なデータはローカルのパソコンでつなぎ合わせるができず、スーパーコンピュータを用いることになるという、隠れた巨大化（目に見えない巨大化）である。そして、3点目は、先の2点目の結果として、シーケンスに必要な人材が多様化したことである。シーケンス作業をする実験助手、基礎解析やデータの視覚化をする分子生物学の知識を持ったシステムエンジニア、データの管理者がまず必要である。ゲノム科学では、実験から高度な情報解析までを必要とする次世代シーケンサーをツールとして研究を行っていくので、ゲノム研究者にもさまざまなバックグラウンドを持つ研究者がいる。例えば、情報のみを扱う研究者、情報解析と実験の両方とも遂行する研究者、実験系の研究者などがある。以下では、その内実を詳細にみていく。

次世代シーケンサーを利用したゲノム科学においては、サンプル調整者の実験系の研究者と情報解析担当者のバイオインフォマティクスの研究者がかかわっている場合が多く、この場合、少なくとも2つの分野の研究

者が存在する。もし、研究テーマがある病態に関する網羅的なゲノム解析ならば、その専門家も研究に参加することになる。例えば、がんをテーマとしたのであれば、がん研究者が入ることが考えられるだろう。こうして様々な分野の研究者や作業員が、数日間にわたる次世代シーケンサーの稼働の都合に合わせて仕事をし、各々の役割をこなしているのが現状である。従来ならば、シーケンスしたいときにサンプル調整の実験を行い、機器に入れ、好きな時に必要な結果を解析すればよかった。しかし、著者が行ったインタビューによると、次世代シーケンサーは、一度の稼働で他のサンプルと一緒にシーケンスすることがあるので、その日程に合わせてサンプル調整を行うか、または次の稼働まで待たなければならぬ。そして、シーケンスには数日間を要し、情報解析担当者が基礎解析まで行ったデータを受け取ることになる。そこから先の解析にも特殊な情報処理の知識と技術が必要だ。次世代シーケンサーは、日本で人材が豊富でない分野の、様々な職種の間を振り回しているとも言えよう。

だが、このようなゲノム科学の巨大化は、たとえばある民族やある病態に特異的な SNP の発見や、オーダーメイド型の医療を推進していくために求められた姿でもある。前述したとおり、サンガーは塩基配列決定法の開発時から、塩基配列に生物にとって必要なすべての情報が含まれていると考え、生物の理解と医学への貢献のために「簡単で速い」方法を目指していた。この記述を踏まえると、ゲノム科学の巨大化は、現象としては国際ヒトゲノム計画から始まったものと考えられるが、塩基配列決定法の開発当初から、医療への貢献を目指しており、巨大化の土台は構築され始めていたと言えるのではない。

ただし、サンガーは“simple and rapid”や“accurate”なシーケンス技術を目指していたのであるが、本論文で明らかにしたとおり、現在の次世代シーケンサーを利用した配列の解読は、簡単でも正確性が高いわけでもない。むしろ複雑化し、読み取りエラーはサンガー法に比べれば高い。サンガーが目指した目標のうち、唯一「高速化」だけが進み、簡単さや正確性の代わりに得たものが大量の配列データなのである。生活習慣が狂い、複雑で、間違いが多い。それでも、研究者は次世代シーケンサーを受容し、集団のゲノム解析をする場合や網羅的な解析を伴うゲノム医学研究に役立てようとしている。はたしてこの機器を制御できていると言えるのかどうかは、まだ評価できる段階にはないだろう。

他方、ゲノム科学は多くの研究参加者の協力があって初めて大規模かつ網羅的な解析を行えるという側面があり、巨大化が進んでいけばいくほど、社会や市民とのかわりが深くなっていく。しかしこれらの巨大化は、

先行する巨大科学の事例と共通するも、専門家ではない一般市民には把握しづらい変化である。特にゲノム科学におけるデータの巨大化は、これまでの巨大科学ではみられない現象であり、ゲノムという目には見えないが、巨大な情報を内包するという特徴に起因する巨大化と考えられる。

つまり、研究者自身もゲノム科学の巨大化の含意をあらためて認識する必要がある。同時に、研究参加者にとっても想像しにくい科学であることを踏まえたうえで、研究の必要性を社会に浸透させる努力が必要ではないだろうか。

6. 終わりに

これまで分析してきたとおり、初期のシーケンス技術は個々の研究者が単独で利用できる範囲の成果に収まっていたが、現在のシーケンス技術は、もはや単独の研究者の手には負えないほど複雑に発展した。それがゲノム配列を高速に大量に解読する「次世代シーケンサー」である。次世代シーケンサーを利用したゲノム解析技術は多様化しており、現在ではゲノム科学を支える主要な機器の一つとなっている。

次世代シーケンサーは、ゲノム解析技術に大きな進展をもたらしたが、本論文で考察してきたとおり、専門家集団のなかでも使いこなせていない事実がある。著しく進展したゲノム解析技術を使いこなし、こうした事態を改善するためには、本論文で分析してきた個別の課題を総合的な課題として捉え直し、早急に取り組む必要がある。

具体的には本論文中で述べた通りであるが、次世代シーケンサーの登場は、ゲノム解析に関わる研究者の研究労働の変化を伴いながらゲノム科学を巨大科学化させた。その変化とは、多くの研究者の分業という研究形態や、多様な分野の知見を結集させる必要性であり、次世代シーケンサーを基軸としたゲノム解析研究が巨大科学化した経緯を示すことが出来たのは本論文の成果である。

また、本論文では、紙幅の関係で、次世代シーケンサーや全ゲノム解析研究に伴う倫理的法的社会的諸課題について言及できていないが、次世代シーケンサーが専門家集団の内外を問わず利用され始め、個人の全ゲノム解析が盛んに行われるようになる前に、検討する必要があるだろう。

参考文献

- 1) Watson, J. D., and Crick, F. H. C. (1953). Molecular structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171, 737-738.
- 2) Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA Sequencing with Chain-Terminating Inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), 5463-5467.
- 3) Maxam, A. M., and Gilbert, W. (1977) A New Method for Sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74(2), 560-564.
- 4) Rabinow, P. (1998) 『PCRの誕生-バイオテクノロジーのエスノグラフィー』(渡部政隆訳) みすず書房 (原著 1996年) .
- 5) Mullis, K. B., and Faloona, F. A. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*, 155, 335-350.
- 6) Galison P., Hevly B. (Eds.), (1992). *Big Science: The Growth of Large-Scale Research*. Stanford University Press.
- 7) Traweek, S. (1998). *Beamtimes and Lifetimes: The World of High Energy Physicists*. Harvard University Press.
- 8) 高木仁三郎(1979)『科学は変わる—巨大科学への批判』社会思想社
- 9) 岸田純之助(1969)『朝日市民教室〈日本と核時代〉2 巨大科学と政治』朝日新聞社
- 10) 平田光司(2002)「『大型科学』論と STS の課題」『科学技術社会論研究』(1), 68-74.
- 11) 井上英二(1990)「ヒト・ゲノム・プロジェクト—医学生物学のビッグサイエンス—」『遺伝』44(5), 40-43.
- 12) 寺内かえで, 田村倫子(2008)「ヒトゲノムプロジェクト—その技術と政治—」『広島工業大学紀要研究編』42, 237-245.
- 13) Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H., and Fischer, D. (1973). Use of DNA Polymerase Primed by a Synthetic Oligonucleotide to Determine a Nucleotide Sequence in Phage f1 DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74(4), 1209-1213.
- 14) Sanger, F., and Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94 (3), 441-446.
- 15) Sanger, F. (2001) The early days of DNA sequences. *Nature Medicine*, 7(3), 267-268.
- 16) 岸宜仁(2004)『ゲノム敗北—知財立国日本が危ない』ダイヤモンド社
- 17) 和田昭允(2005)『物理学は越境する ゲノムへの道』岩波書店.
- 18) 服部正平(2003)「ヒトゲノム配列の完全解読」『科学』73(4), 393-399.
- 19) 榊佳之(1990)「遺伝子の構造解析」『遺伝子工学実験』48-55.
- 20) 村松正実[編](1990)『ラボマニュアル遺伝学 (増補版)』丸善出版.
- 21) 中山広樹, 西方敬人(1995)『細胞工学別冊 目で見える実験ノートシリーズ バイオ実験イラストレイテッド ②遺伝子解析の基礎』秀潤社
- 22) 小笠原直毅, 高見英人, 久原哲, 服部正平(2001)『生物化学実験法 46 ホールゲノムショットガン法によるゲノム解析とアノテーション』学会出版センター.
- 23) 金井昭教, 菅野純夫, 鈴木穰(2008)「ChIP-Seq」牛島俊和・眞貝洋一編『実験医学別冊 注目のバイオ実験シリーズ エピジェネティクス実験プロトコール DNAメチル化とヒストン修飾を網羅的・領域特異的, 定量的に解析する実験手法のすべて』(pp. 247-260.) 羊土社
- 24) 菅野純夫(2008)「次世代シーケンサーのインパクト」『医学のあゆみ』225(9), 753-757.
- 25) 水島-菅野純子(2001)『ゲノムでわかることできること—注目のゲノム医学の基礎知識から最新のゲノム情報まで』羊土社.
- 26) DOE Genomics Timeline
[<http://www.genomicscience.energy.gov/program/timeline.shtml>]
2013年9月1日アクセス]
- 27) 榊佳之(2007)『ゲノムサイエンス』講談社
- 28) Cold Spring Harbor Oral History Collection (2003). *James D. Watson on The 1986 Symposium: The Molecular Biology of Homo Sapiens*.
[<http://library.cshl.edu/oralhistory/interview/cshl/symposia/1986-symposium-homo-sapiens/>]
2013年9月1日アクセス]
- 29) Dulbecco, R. (1986) A Turning Point in Cancer Research: Sequencing the Human Genome. *Science*, 231, 1055-1056.
- 30) Watson, J. D., and Berry, A. (2003) 『DNA すべてはここから始まった』(青木薫訳) 講談社 (原著 2003年) .
- 31) 松原謙一, 中村桂子(1996)『ゲノムを読む—人間を知るために』紀伊国屋書店
- 32) 宋碩林, 一戸 敦子, 菅野純夫(2010)「シーケンシング技術開発の歴史といま、そして未来」『実験医学』28(9), 1442-1448.
- 33) 鈴木穰, 菅野純夫(2008)「次世代シーケンサーを用いたゲノム情報の解読」『生物の科学 遺伝』62(2)
- 34) Mardis, E. R..(2013) Next-Generation Sequencing Platforms. *Annual Review of Analytical Chemistry*, 6, 287-303.
- 35) 近藤直人(2012)「NGS 購入ガイド (必要なもの)」菅野純夫・鈴木穰編『次世代シーケンサー目的別アドバンスメソッド』学研メディカル秀潤社
- 36) 森下真一(2009)「次世代高速シーケンサーの特性と情報処理」『蛋白質核酸酵素』54(10), 1239-1247.
- 37) 内閣官房健康・医療推進本部(2014)『医療分野の研究開発

に関する総合戦略』(医療分野の研究開発に関する専門調査会報告書)。

- 38) 高木仁三郎(1981)『プルトニウムの恐怖』岩波書店。
39) 日本放送協会海外取材班(1969)『巨大科学』日本放送協会。

謝辞

東京大学医科学研究所公共政策研究分野の所属メンバーに感謝の意を表す。本研究は、科研費(特別研究員奨励費)「次世代型ゲノム解析技術の開発・応用と倫理的・法的社会的課題に関する研究」の助成を受けている。

-
- i) 本論文では、塩基配列を実験またはシーケンサーのような解析機器を通して知ることを「解読」とした。
ii) 従来とは異なる原理で大量の塩基配列を解読する機器のことを指す。シーケンサーを利用している研究者の間において呼ばれている名称、論文のタイトル等にも用いられ、広く使われている。
iii) 塩基配列を知ることが「解読」としたことに対し、エラーを修正し、利用できるデータとしての配列情報にすることを「決定」とした。
iv) 単位時間あたりに解読できる塩基配列数をスループットといい、解読できる塩基数が多いほどハイスループットであると表現する。

DEVELOPMENT AND CHALLENGES OF GENOME ANALYSIS TECHNOLOGY -DNA SEQUENCING TECHNOLOGY BECAME LARGE-SCALE TECHNOLOGY-

Takako ARAUCHI¹, Yusuke INOUE², Taichi ISOBE³, and Kaori MUTO⁴

¹Master of Bioscience, Doctor Course Student, Medical Genome Science, Graduate School of Frontier Science, The University of Tokyo (E-mail: arauchi@hgc.jp)

² Ph.D. (Public Health) Assistant Professor, Department of Public Policy, The Institute of Medical Science, The University of Tokyo (E-mail: yinoue-kyt@umin.ac.jp)

³M.A. (Interdisciplinary Information Studies) JSPS Postdoctoral Research Fellow, Department of Public Policy, The Institute of Medical Science, The University of Tokyo (E-mail: tisobe@ims.u-tokyo.ac.jp)

⁴ Ph.D. (International Community Health) Professor, Department of Public Policy, The Institute of Medical Science, The University of Tokyo (E-mail: krmt@ims.u-tokyo.ac.jp)

DNA sequencing technology was developed because of the advancements in the Human Genome Project (1990–2003). After that, a “next-generation sequencer,” which produces thousands or millions of sequences, was introduced. In the field of life sciences and medicine, the development of sequencing technology has changed the target of research from one gene to the complete genome. The aim of this study was to show the features of the next-generation sequencing technology and to examine the necessary environment required for operating it. Moreover, genome research with many challenges was discussed in terms of “big science” which rely on large-scale project usually funded by the national government and are required the cooperation of many various experts.

Key Words: *DNA Sequencing technology, next-generation sequencer, genome research, the Human Genome Project, big science*