

Human Glycome Project への世界と日本の動き ～本特集の序に代えて～

門松 健治

ヒトゲノムが解読され、2万余種のタンパク質しか体に存在しないと知ったとき、タンパク質や脂質に無限に近い多様性を与え、自らも機能を担う糖鎖を無視できないことに、研究者たちは気づき始めている。この中で、ヒトの糖鎖の網羅的解析が進もうとしている。4次元の変化を見せる糖鎖のダイナミクスは生理的あるいは病的事象に最も近いところにあると思われる。両者を結びつけるには、対象や手法の選択を含む研究デザインを十分に練る必要がある。さらに糖鎖構造決定のためのハイスループット分析技術と糖鎖・糖ペプチド同定のためのインフォマティクス技術の開発が必須である。ヒトゲノム計画を好例に、データシェアリングを基本精神として他のオミクスデータなどとの組み合わせ解析が可能になれば、生命についての理解は新しいフェーズに入ることになるだろう。

1. 糖鎖と生命科学

核酸（ヌクレオチドの鎖）、タンパク質（アミノ酸の鎖）と糖鎖（糖の鎖）。この三つは生体内ポリマーの代表であり、三大生命鎖と呼ぶにふさわしい働きをしている。しかるに、核酸、タンパク質に比べて、糖鎖の認知度は高くなく、むしろその構造と機能の多様性、複雑性のために研究者からは避けられてきたきらいがある。だが、ヒトゲノムが解読され、2万余の限られた種類のタンパク質しかヒトの体に存在しないと知ったとき、タンパク質や脂質に無限に近い多様性を与え、自らも機能を担う糖鎖を無視できないことに、研究者たちは気づき始めている。このような動きの中で、ヒトの体の糖鎖の網羅的解析が進もうとしている。

糖鎖の前に、少しだけゲノムを振り返ってみたい。1990

年にHuman Genome Project（HGP；ヒトゲノム計画）が米国のエネルギー省などの支援によって始まった。日本や欧州も参加した一大プロジェクトであった。当初は15年の計画であったが、シーケンサーならびにインフォマティクスの進歩により、2000年にはゲノムドラフトが完成した（表1）。これは予定よりも2年早い完成であった。この加速を後押ししたのは、民間企業との「競争」であった。John Craig Venterはインフルエンザ菌の全ゲノムを世界で初めて解読したが、生物の全ゲノム情報の販売を目的に設立されたCelera Genomics社社長に就き、ヒトゲノム解読に参入した。Celera社が用いたショットガン・シーケンス法は高速の配列解読を可能とし、HGPもこの方法を導入した。これにより全ゲノム解読は加速した。Celera社は当初解読配列の秘匿を方針としていたが、科学の進歩を妨げるものとして批判を浴び、1996年にバミューダで開かれた会議においてオープンアクセスポリシー（バミューダ原則）の合意に至った（表1）。こうして、2000年にゲノムドラフトが発表され、2001年にはHGP側は*Nature*誌¹⁾、Celera社は*Science*誌²⁾に、その配列に対する分析ならびにドラフト構築の手法を発表した（表1）。

日本の生物物理学者、和田昭允はHGP開始から遡ること約10年の1981年に遺伝子配列の高速自動解析装置を開発するプロジェクトを開始していた。富士写真フイルム、エプソンなどの企業が参加し、3年後には自動解析装置の試作機もできた。しかし官庁や研究者からの積極的な支

名古屋大学大学院医学系研究科生物化学講座（〒466-8550 名古屋市中区鶴舞町65）

Introduction: Human Glycome : Trends in Japan and World

Kenji Kadomatsu (Department of Biochemistry, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan)

本論文の図版はモノクロ（冊子版）およびカラー（電子版）で掲載。

DOI: 10.14952/SEIKAGAKU.2020.920303

© 2020 公益社団法人日本生化学会

表1 歴史的出来事

糖鎖と生命科学

- 2012年 National Academy of Scienceの諮問機関が“*Transforming Glycoscience*”をレポート
 2014年 JCGG (Japan Consortium for Glycobiology & Glycotechnology) が『*Glycoscience : Biology and Medicine*』を出版
 2016年 AMEDとNIH連携覚書締結「NIHは、日本が世界に先駆けている糖鎖研究の分野での連携に強い関心を持っている」
 2018年 日 (GlyCosmos) 米 (GlyGen) 欧 (Glycomics@ExPASy) の糖鎖インフォマティクスが連合してGlySpace Allianceを開始
 2019年 JCGGが『*Glycoscience : Basic Science to Applications*』を出版

Human Genome Project

- 1990年 Human Genome Project開始
 1996年 バミューダ原則締結
 2000年 ゲノムドラフト発表
 2001年 Human Genome Projectが*Nature*誌に、Celera Genomics社が*Science*誌に分析結果を発表

Human Glycome Project, Human Glycoproteomics

- 2004年 HUPO (Human Proteome Organization) がHuman Glyco/Proteomics Initiativeを提唱
 2017年 HUPOがHuman Glycoproteomics Initiativeを開始
 2018年 Human Glycome Project (欧米) 開始
 2019年 Human Glycome Project (日本) をマスタープラン2020に提案

援をその後得ることはかなわなかった。1980年代前半は、日本が米国に急迫してきた時代でもある。家電や自動車の二の舞を遺伝子でも味わうのか。このような時代の背景が、米国発のHGPの実現にはあったのかもしれない³⁾。

さて、糖鎖である。核酸やタンパク質の研究が長足の進歩を遂げた現代において、もう一つの生命鎖である糖鎖の科学を進展させるにはどうしたらいいのか。2012年に米国National Academy of Scienceが諮問した委員会が“*Transforming Glycoscience*”と題したレポートを発表した。ここには10年後あるいはそれ以降の糖鎖科学の向かうべき姿と具体的なロードマップが示された(表1)。日本では2014年に日本糖鎖科学コンソーシアム(Japan Consortium for Glycobiology & Glycotechnology : JCGG)が『*Glycoscience : Biology and Medicine*』を出版した。2016年には日本医療研究開発機構(AMED)と米国National Institute of Health (NIH)が連携の覚書を交わしたが、その中で「NIHは、日本が世界に先駆けている糖鎖研究の分野での連携に強い関心を持っている」とうたった(表1)。糖鎖インフォマティクスの国際連携も本格化し、2018年に、日本のGlyCosmos、米国のGlyGen、欧州のGlycomics@ExPASyが連合して糖鎖インフォマティクスの国際共同体GlySpace Allianceを設立した。さらに2019年JCGGは『*Glycoscience : Basic Science to Applications*』を出版し、日本と世界が協調し新たな糖鎖科学の時代を築くべく指針とロードマップを提示した(表1)。

こうしてみると、この10年の間に糖鎖科学の動きが急となっていることに気づく。予想をはるかに下回る2万余りに限られたヒトの遺伝子数はイネ科植物より少なく、ウニと大差がない。non-coding RNAなどの種類と機能が爆発的に増えて再注目されており、これを含めてヒトと他の

動物を差別化するいくつかの観点が提示されているが、遺伝子にコードされるタンパク質は2万余の種類という事実には変わりはない。そうすると翻訳後のタンパク質のダイナミクスがヒトという高等生物を作り生かしているのだらうと予想される。そのダイナミクスの一翼を担うのが翻訳後修飾といわれる一連の反応であろう。リン酸化、アセチル化、ユビキチン化、パルミトイル化など多種の翻訳後修飾がタンパク質には付与されるが、中でも糖鎖が与える多様性は無尽蔵に近い。なぜなら、糖鎖は一つの形にとどまるのではなく、結合様式、鎖長も多種にわたり、しかも同じタンパク質の同じ箇所につく糖鎖でさえ構造が異なる可能性が高い。それほど多様である。そして、時代はやっと追いついてきた感がある。すなわち、網羅的糖鎖情報の取得が可能になった時代に入ってきた。まず、糖鎖構造解析技術が進歩して糖鎖構造の多くが解けるようになってきた。そして、インフォマティクスの進歩により糖鎖のアノテーション(ある構造を持つ糖鎖が付加したペプチド断片の同定)も可能となってきた。

ヒトの体の場所と時間の軸で一網打尽に糖鎖の構造を決めてみたい。この夢がかなえば、我々の生命に関する理解は飛躍的に深まるであろう。この糖鎖情報がゲノムやタンパク質あるいは脂質などの情報と組み合わせると、これまでの常識が覆される場面に何度も出会うことになるだろう。このような背景からHuman Glycome Projectが始まろうとしている。

ここで混乱を避けるためにGlycome (グライコム)、Glycoproteome (グライコプロテオーム)、Glycomics (グライコミクス)、Glycoproteomics (グライコプロテオミクス)の言葉の定義をしておきたい。前二者は分子の総体であり、後二者はそれらの知識体系や研究分野を指す。Gly-

come/GlycomicsとGlycoproteome/Glycoproteomicsを厳密に区別すると、前者は糖鎖をタンパク質などから切り離し糖鎖単体として構造を決定するものである。一方、後者はタンパク質の糖鎖付加部位とそれぞれの糖鎖構造を決定する。当然、Glycoproteome/Glycoproteomicsにはより多くの労力と技術を要することになるが、糖タンパク質についてより正確な情報が得られるし、生理的あるいは病的状態の把握に役立つ可能性が高い。一方、Glycome/Glycomicsは、糖タンパク質の糖鎖に限らず、スフィンゴ糖脂質の糖鎖、遊離オリゴ糖をも対象にする。特に糖タンパク質糖鎖のうち、N型糖鎖に関してはPNGase Fという網羅的にN型糖鎖をタンパク質から切り離すことのできる酵素があるため、多数のサンプルを解析することが可能であり、総体としての糖鎖構造の変化が生理的・病的状態の把握にクリティカルな場合は大きな威力を発揮するといえる。

2. Human Glycome Projectへの挑戦

1) 世界と日本の動き

2004年に国際組織Human Proteome Organization (HUPO)からHuman Glyco/Proteomics Initiativeが提示され、糖タンパク質の解析の重要性が提唱された。この考えは2017年から始まったHUPOのHuman Glycoproteomics Initiative (HGI)に引き継がれた(表1)。2017年からのHUPOのHGIはオーストラリアMacquarie大学のNicki PackerとMorten Thaysen-Andersenによってリードされている。上述のとおりGlycoproteome/Glycoproteomicsには分析技術の進展が欠かせないものであるが、同時にまたインフォマティクスの技術、つまり、出てきたデータがどの糖鎖構造を指し、どのタンパク質部位に付加したものといえるか、というアノテーション技術が欠かせない。HGIでは現在、特に後者にフォーカスしている。研究は、アノテーションのソフトウェアを開発するdevelopers(9チーム)とそのソフトウェアを使って実際のデータを解析してアノテーション(ある構造を持つ糖鎖が付加したペプチド断片の同定)を行うusers(13チーム)からなる。ヒト血清を材料にLC-MS/MSによって得られた二つの糖ペプチドデータセットをもとに、developersが開発したソフトウェアを使ってusersがN型およびO型の糖ペプチドを同定する。これらの結果をHGIの委員会が評価する。その結果、(半)自動で糖ペプチドを同定するソフトウェアは数多く存在していること、一方でこれらによる同定結果は必ずしも一致しないこと、が判明した。したがって、現状ではGlycoproteomicsには分析技術のみならずインフォマティクス技術にもいまだ大きな課題が残っており⁴⁾、これも視野に入れながらさらなる開発が進むと思われる⁵⁻⁸⁾。

話題をHuman Glycome Projectに移そう。2018年、Human Glycome ProjectがHarvard大学のRick Cummingsと、クロアチアZagreb大学のGordan Laucをリーダーとしてスタートした。特に、LaucはGenosという糖鎖分析を専門と

する会社を設立しており、糖鎖分析を担う。重要なのは2019年6月にNew England Biolabs (NEB)とWatersの2企業が本事業との提携を決めたことである。これにより本事業は加速する可能性がある。すなわち、NEBは高速にN型糖鎖を切り出す酵素Rapid PNGase Fを作っており、Watersはそれを用いて質量分析器までの処理キットGlycoWorks RapiFluor-MS N-Glycan Kitを販売している。これらをHuman Glycome Projectに無償で提供し、Genosがハイスループットで糖鎖分析を行うものである。これによって年間1万検体、3年間で3万検体の分析が可能となるという⁹⁾。

日本では2019年に日本学術会議によるマスタープラン2020の公募に際して「ヒューマングライコムプロジェクト(Human Glycome Project)」が大型研究計画として提案された。この提案では健康高齢者コホートと認知症コホートを軸に、Glycomeを解読し、ゲノムや画像などの他の情報との組み合わせにより、生理的老化や疾患の真の理解につなげようとする。一方で、解析技術の開発にも注力する。もう一つ大事なのは、ヘッドクォーターを中心に糖鎖解析の技術教育やワンストップ相談窓口を設けることである。これらの提案された取り組みは、糖鎖科学で世界をリードしてきた日本の強みを生かせるものであり、また、自然科学の中で発展が最も期待される分野の一つである糖鎖科学のすそ野を広げる意味で大きな一歩となるはずである。中でもGlycomeと他の情報との組み合わせ解析の例はいまだ少なく、日本が新しい分野をリードできる可能性がある。

2) 現状と課題

Glycome解読にあたっては、糖鎖がタンパク質から切り出されなければならない。N型糖鎖はPNGase Fがほぼ万遍なく切り出すのでよいが、O型糖鎖についてはいまだそのような酵素は見つかっていない。また、これらの酵素などによる前処理は短ければ短いほど分析の効率化につながるが、現状はこの点においてもいまだ大きな伸び代を残している。その他、このあたりの分析技術の課題については本特集の古川潤一の稿を参考にされたい。また、インフォマティクスの技術開発も引き続き重要である(本特集の木下聖子の稿を参考にされたい)。一方、Glycoproteomeの解読が高速でできる時代になることが望ましい。既述のとおり、解析技術に加えてインフォマティクス技術も開発が進行中である⁵⁻⁸⁾。また、糖鎖情報はタンパク質情報と組み合わせることでより精緻でかつ意味のある情報となる。この意味で組み合わせデータベースの構築はとても重要であり、日本ではJST/NBDCが支援するライフサイエンスデータベース統合化推進事業のGlyCosmos Portal、米国ではNIHが支援するGlyGen Project、そしてスイスバイオインフォマティクス研究所のGlycomics@EXPASYが糖タンパク質やその他関連データを統合し、ウェブ上で公開している。またこの3か所が合意し、GlySpace Allianceという共同体としてデータをすべて共有することになっている。

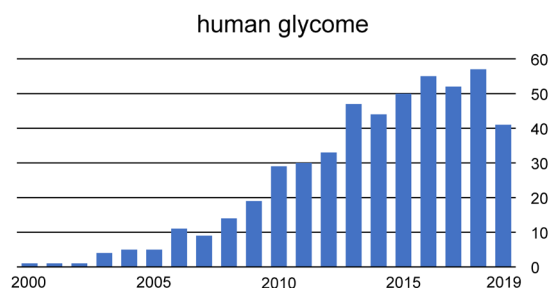


図1 Human Glycomeの論文数推移

ゲノムと違い、GlycomeやGlycoproteomeは場所と時間で変化する。生理的あるいは病的状況で4次元のダイナミクスを示すことになる。これこそがGlycomeやGlycoproteomeの醍醐味であり、表現型の表出に最も近いともいえる。したがって、マスタープラン2020へ提案したHuman Glycome Projectの老化と認知症に限らず、対象とすべき疾患は無限である。たとえば、抗生物質の使用によって起こる代表的な腸内感染症、*Clostridioides difficile*感染に対する便移植治療の前後での患者血清N型糖鎖構造の変化は面白い¹⁰⁾。便移植治療成功例ではN型糖鎖の複雑性の減少がみられる。つまり、高度分岐してガラクトースが4個あるいはシアル酸が3個つくようなN型糖鎖の量が減り、分岐の少ないN型糖鎖は相対的に増える。

3. おわりに

PubMedでHuman Glycomeをキーワードで調べると図1に示すような論文数の推移である。いまだ少数の論文しか出ていないことがわかる。しかもここには網羅的な解析は少ししか含まれておらず、一方で、技術開発に関する論文も多くみられる。つまり、Human Glycome, Human Glyco-

proteomeの網羅的解読は緒についたばかりといえる。換言すると今がチャンスである。網羅的解読は論文数を爆発的に増やすだろうが、何よりも真の意味で我々を生命の理解に近づける。

文 献

- 1) International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
- 2) Craig Venter, J.C., et al. (2001). *Science*, **291**, 1304–1351.
- 3) 瀬川至朗 (2017) 米国エネルギー省とヒトゲノム計画、冷戦後の科学技術政策の変容：科学技術に関する調査プロジェクト報告書、国立国会図書館。
- 4) <https://www.hupo.org/HUPOST/7192013>
- 5) Thaysen-Andersen, M., Packer, N.H., & Schulz, B.L. (2016) Maturing Glycoproteomics Technologies Provide Unique Structural Insights into the N-glycoproteome and Its Regulation in Health and Disease. *Mol. Cell. Proteomics*, **15**, 1773–1790.
- 6) Hu, H., Khatri, K., & Zaia, J. (2017) Algorithms and design strategies towards automated glycoproteomics analysis. *Mass Spectrom. Rev.*, **36**, 475–498.
- 7) Zeng, W.F. (2017). *Nat. Commun.*, **8**, 438.
- 8) Narimatsu, H., Kaji, H., Vakhrushev, S.Y., Clausen, H., Zhang, H., Noro, E., Togayachi, A., Nagai-Okatani, C., Kuno, A., Zou, X., et al. (2018) Current Technologies for Complex Glycoproteomics and Their Applications to Biology/Disease-Driven Glycoproteomics. *J. Proteome Res.*, **17**, 4097–4112.
- 9) <https://bioengineeringcommunity.nature.com/users/59087-vivien-marx/posts/50386-sugar-rush>
- 10) Monaghan, T.M., Pučić-Baković, M., Vučković, F., Lee, C., Kao, D., Wójcik, I., Kliček, F., Polyarchou, C., Roach, B., Louie, T., et al.; Human Glycome Project. (2019) Decreased Complexity of Serum N-glycan Structures Associates with Successful Fecal Microbiota Transplantation for Recurrent *Clostridioides difficile* Infection. *Gastroenterology*, **157**, 1676–1678.e3.