

# マルチオミクス

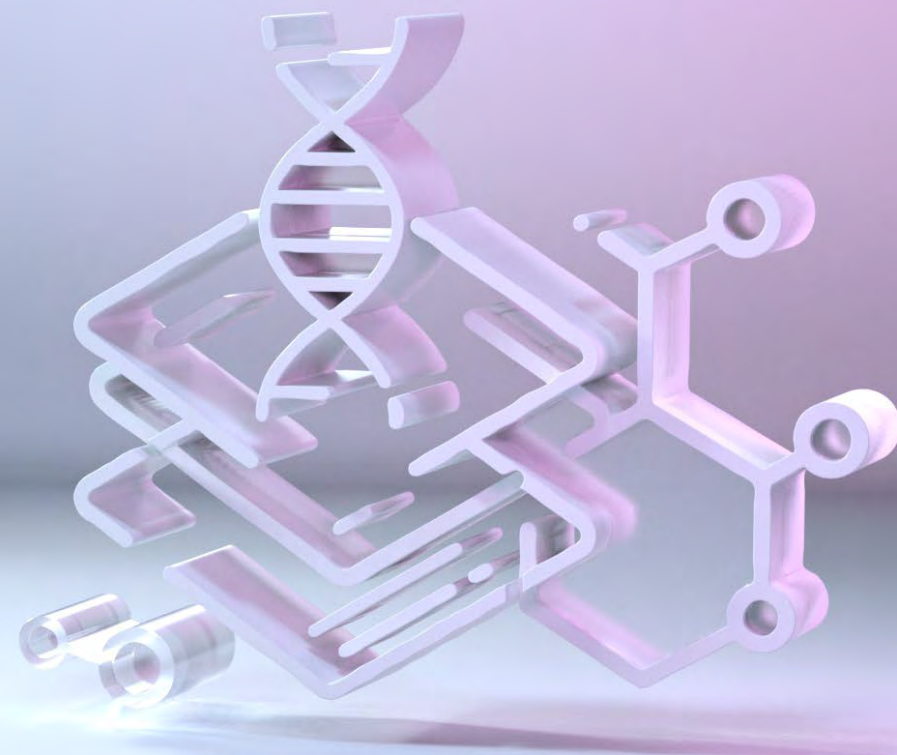
## 分子生物学の未来

リサーチ担当者: Simon Barnett (ARK Invest アナリスト)

ゲノム、トランスクリプトーム、プロテオームなどの「オミクス(生体中に存在する分子の研究)」を活用することで、科学者たちは画期的な発見を臨床的なソリューションに結びつけることができるようになるでしょう。特にハイスループット・プロテオミクスに焦点を当てた新しい分析・ソフトウェアツールによって、科学者たちは生命、疾病、健康の体系を解き明かすことができるようになるはずです。

ARKのリサーチによると、生命科学分野、基礎研究、トランスレーショナルリサーチ、人々の健康増進、分子診断などを含むマルチオミクスの売上は、今後5年間にわたり年率22%のペースで増加し、約1,100億ドルから3,000億ドルに拡大する見通しです。

上記の予測は、限定的なものであり、その信頼性を保証するものではありません。  
投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。





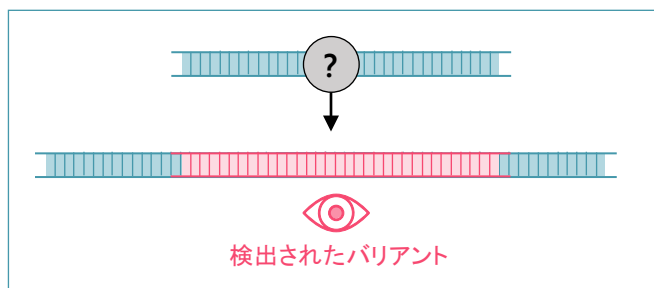
# 「ヒトゲノム計画」は2021年にようやく終了

次世代シーケンサー(NGS)を用いて作成された2003年のヒトゲノムドラフト配列は不完全なものでした。Telomere-to-Telomere (T2T) Consortiumは、ロングリードシーケンス(LRS)法を組み合わせ使用し、その過程でさらに2億塩基対(bps)のDNAと1,500の遺伝子を新たに発見し、完全なヒトゲノム配列を発表しました<sup>1</sup>。ARKでは、この歴史的に価値の高い成果が、これまで解明されていなかった生物学を解明する新しい生命科学のツールや方法を導くことになるとみています。

LRSは、白黒の画像に色をつけるように、ゲノムの見え方をこれまでになく鮮明にします。

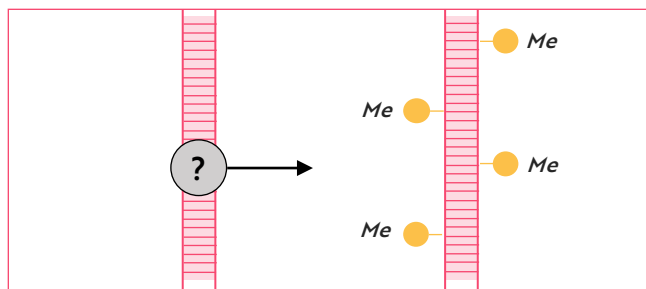
## 構造上の変異

LRSは、SRSに比べて少なくとも2倍以上の構造上の変異をもつことが特徴です<sup>2</sup>。



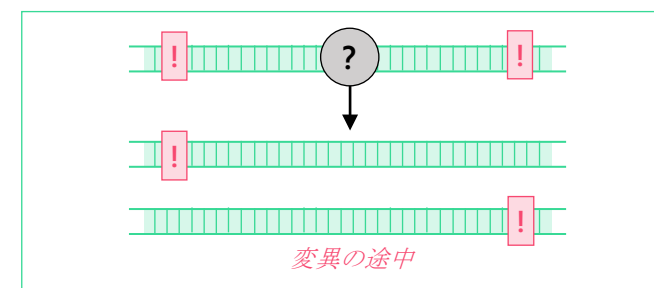
## エピジェネティクス

LRSは、メチル化などのエピジェネティック情報を、別途サンプルを用意することなくネイティブに取り込むことができます<sup>3,4</sup>。



## ロングレンジ・フェージング

LRSは2倍体のゲノムを識別し、突然変異が「母方」にあるのか「父方」にあるのかを確認することを可能にします<sup>5</sup>。



上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] Nurk, Sergey, et al. "The Complete Sequence of a Human Genome." *BioRxiv*, Cold Spring Harbor Laboratory, 1 Jan. 2021, <https://www.biorxiv.org/content/10.1101/2021.05.26.445798v1>.

[2] "Epigenetics and Methylation Analysis." *Oxford Nanopore Technologies*, 26 Oct. 2021, <https://nanoporetech.com/applications/investigation/epigenetics-and-methylation-analysis>.

[3] Ebert, Peter, et al. "De Novo Assembly of 64 Haplotype-Resolved Human Genomes of Diverse Ancestry and Integrated Analysis ..." *BioRxiv*, Cold Spring Harbor Laboratory, 1 Jan. 2020, <https://www.biorxiv.org/content/10.1101/2020.12.16.423102v1>.

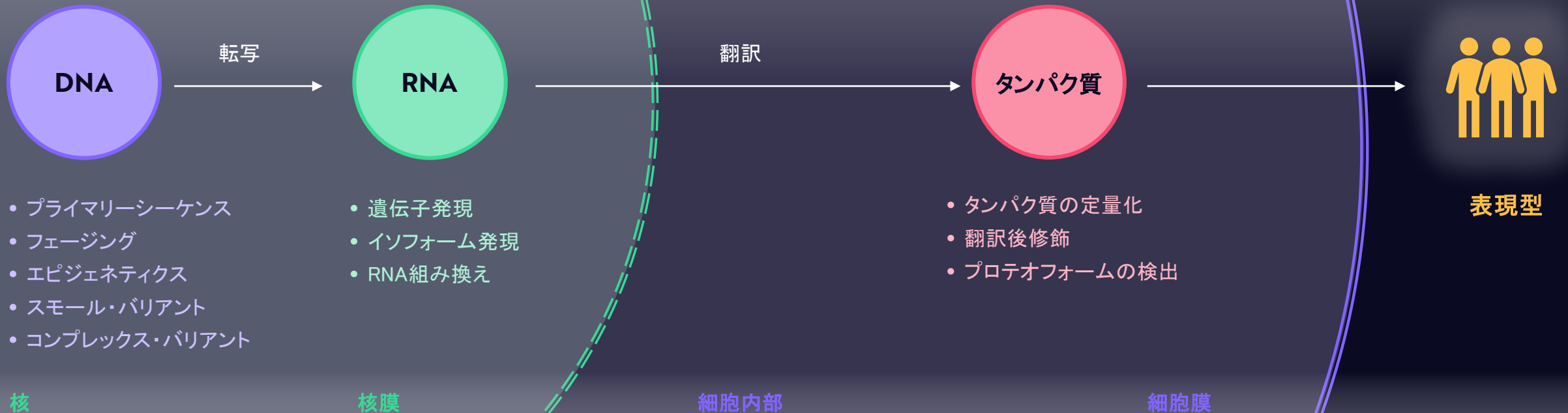
[4] "Direct Detection of DNA Methylation during Single-Molecule, Real-Time Sequencing." *PacBio*, 2010, <https://www.pacb.com/publications/direct-detection-of-dna-methylation-during-single-molecule-real-time-sequencing/>.

[5] Botton, Mariana R., et al. "Phased Haplotype Resolution of the SLC6A4 Promoter Using Long-Read Single Molecule Real-Time Sequencing." *MDPI*, Multidisciplinary Digital Publishing Institute, 12 Nov. 2020, <https://www.mdpi.com/2073-4425/11/11/1333>.



# セントラルドグマ は生物の中における情報の伝達の仕組みを解析する概念

要約すると、セントラルドグマでは、**DNA** (ゲノム) が **RNA** (トランスクリプトーム) に転写され、それが最終的にタンパク質 (プロテオーム) に翻訳されるとしています<sup>1</sup>。タンパク質は、生命維持に必要なほとんどすべての機能を担っていますが、その機能が変化すると病気の原因となります。セントラルドグマの各過程の間の相互作用を理解することで、予測、診断、そして基本的な生物学的洞察の飛躍を実現する能力が向上します。



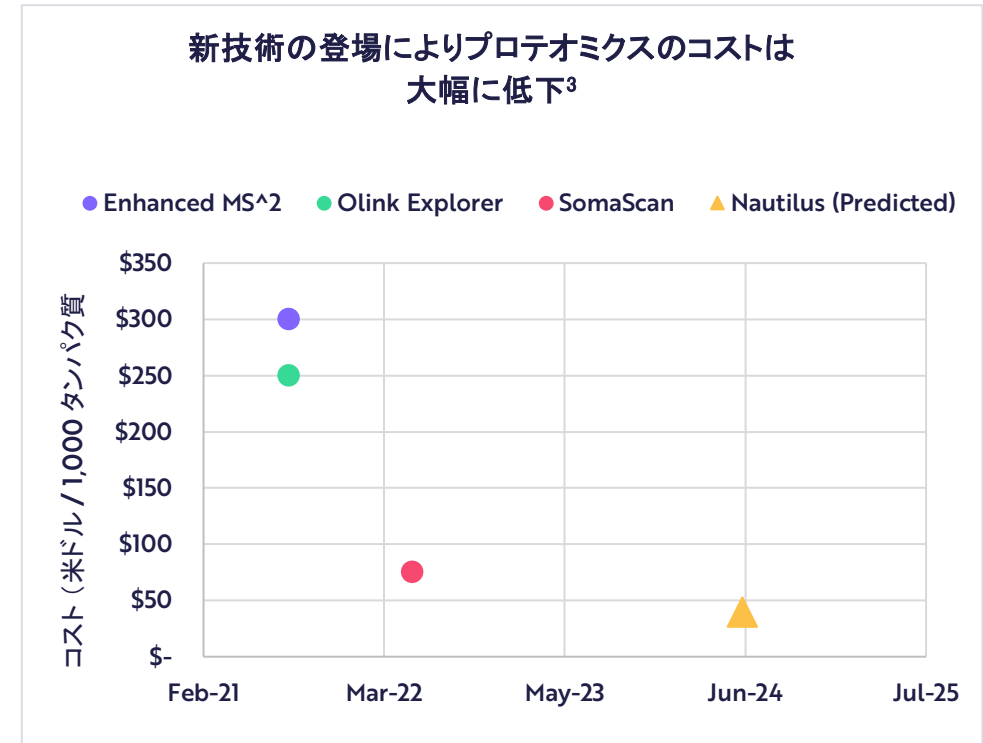
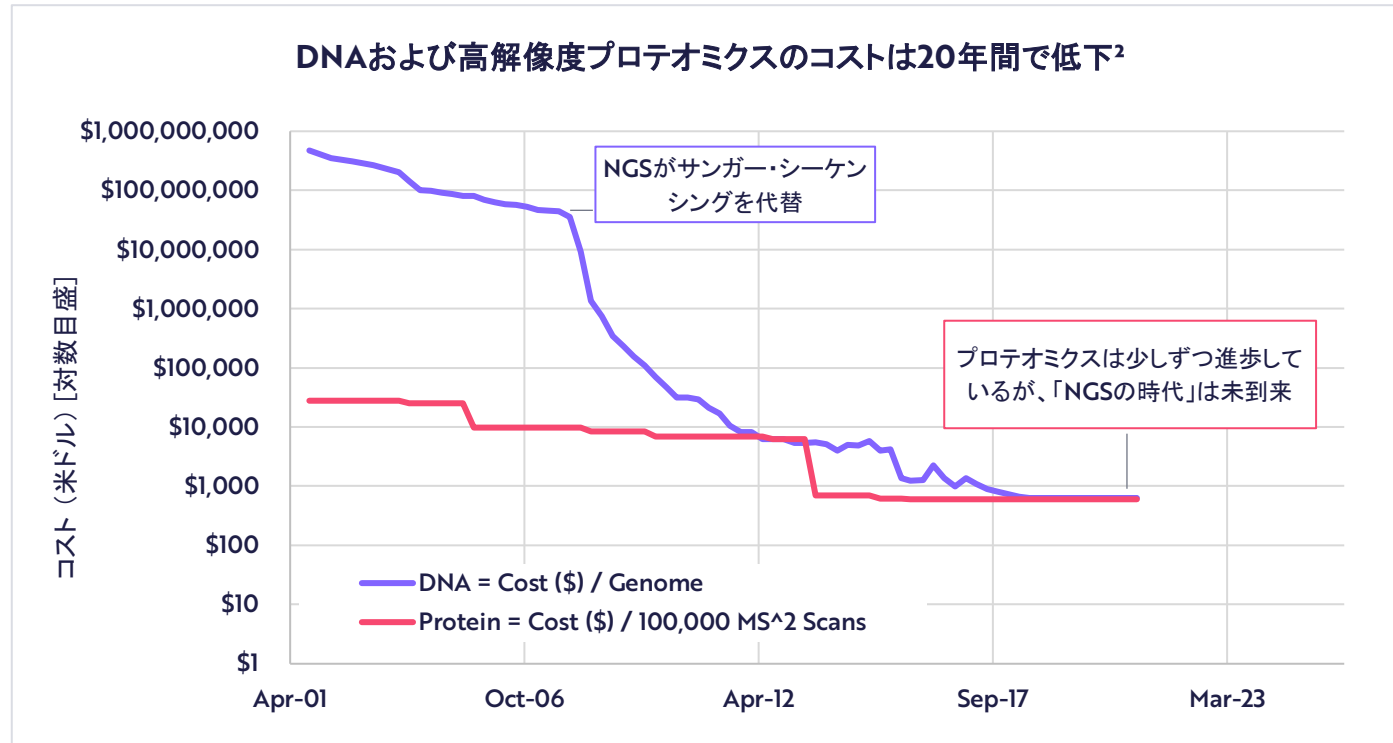
上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] Malone, Cindy S. "Central Dogma of Genetics - CSUN.EDU." *csun.edu*, <https://www.csun.edu/~cmalone/pdf360/Ch13-1transcription.pdf>.



## 高額な費用等、様々な課題を抱える包括的臨床プロテオミクス

血清(プラズマ)は、複雑なヒトプロテオームを偏りなく見ることができる唯一のサンプルの種類です。ARKでは、臨床プロテオミクスが人間の健康の指標として非常に大きな可能性を秘めており、セントラルドグマの概念を理解する鍵になると考えています<sup>1</sup>。



MS<sup>2</sup> = "Tandem Mass Spectrometry (タンデム式質量分析装置)"の略で、血漿のような複雑なプロテオミクスサンプルを分析するための一般的な手法です。

上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] Blume, John E., et al. "Rapid, Deep and Precise Profiling of the Plasma Proteome with Multi-Nanoparticle Protein Corona." *Nature News*, Nature Publishing Group, 22 July 2020, <https://www.nature.com/articles/s41467-020-17033-7>.

[2] ARK Investment Management LLC, 2021, Data sourced from the National Human Genome Research Institute (NHGRI) and Nature Reviews, Chemistry

[3] ARK Investment Management LLC, 2021, Data sourced from Seer (SEER), Olink Proteomics (OLK), SomaLogic (SLGC), and Nautilus (NAUT)

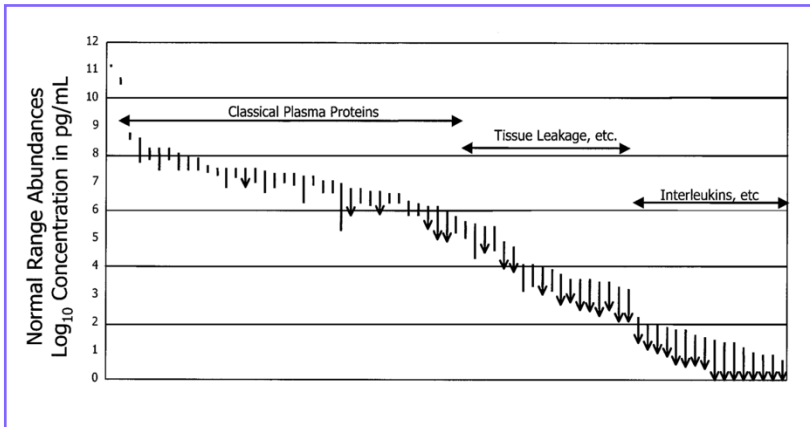


## ルネッサンス期にある質量分析(MS)

質量分析(MS)は、タンパク質の質量や電荷の違いを利用して、複雑なプロテオミクスサンプルを一度に分析する技術です。MSは、血漿中のタンパク質レベルを検出できないことが多い手作業による分析を必要とするため、規模の拡大が難しいとされてきました<sup>1</sup>。

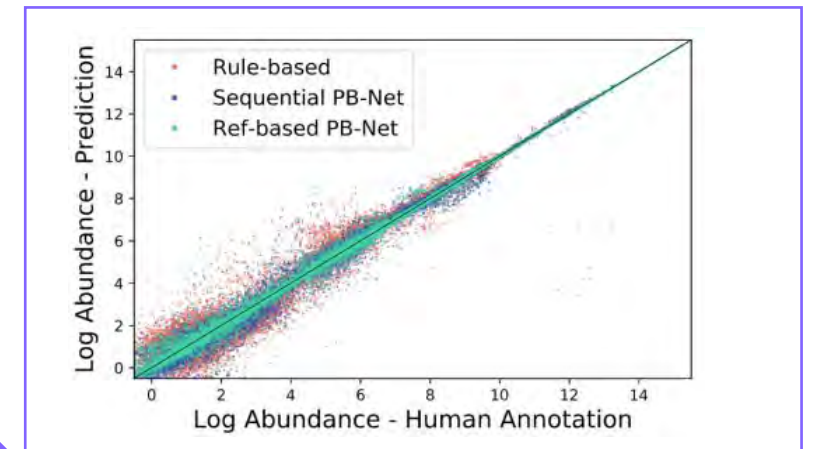
MSの欠点を補う方法として新しいサンプルの前処理と分析を行なうようになりました。MSは一般的な技術であり、市場はすぐにこれらの進歩を取り入れることができるでしょう。

調整可能なナノ粒子のような新しいサンプル前処理法は、血漿プロテオームのダイナミックレンジを圧縮し、低濃度および高濃度のタンパク質を検出しやすくすることができます<sup>1,2</sup>。



- 新しいMS装置では、**Scanning SWATH**のような巧妙な方法により多くのデータを入手できます<sup>3</sup>
- **Scanning SWATH**は、血漿タンパク質の検出率を70% (CV: 6.4%) 向上させ、サンプルの前処理時間を5分に短縮します

**OpenPIP**のようなディープニューラルネットワーク(NN)ツールは、MSのスペクトルデータの解析に必要な時間を短縮します<sup>4</sup>。



上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] Blume, John E., et al. "Rapid, Deep and Precise Profiling of the Plasma Proteome with Multi-Nanoparticle Protein Corona." *Nature News*, Nature Publishing Group, 22 July 2020, <https://www.nature.com/articles/s41467-020-17033-7>.

[2] Anderson, Norman, and N. Leigh Anderson. "The Human Plasma Proteome." *Molecular and Cellular Proteomics*, [https://www.mcponline.org/article/S1535-9476\(20\)34246-8/fulltext](https://www.mcponline.org/article/S1535-9476(20)34246-8/fulltext).

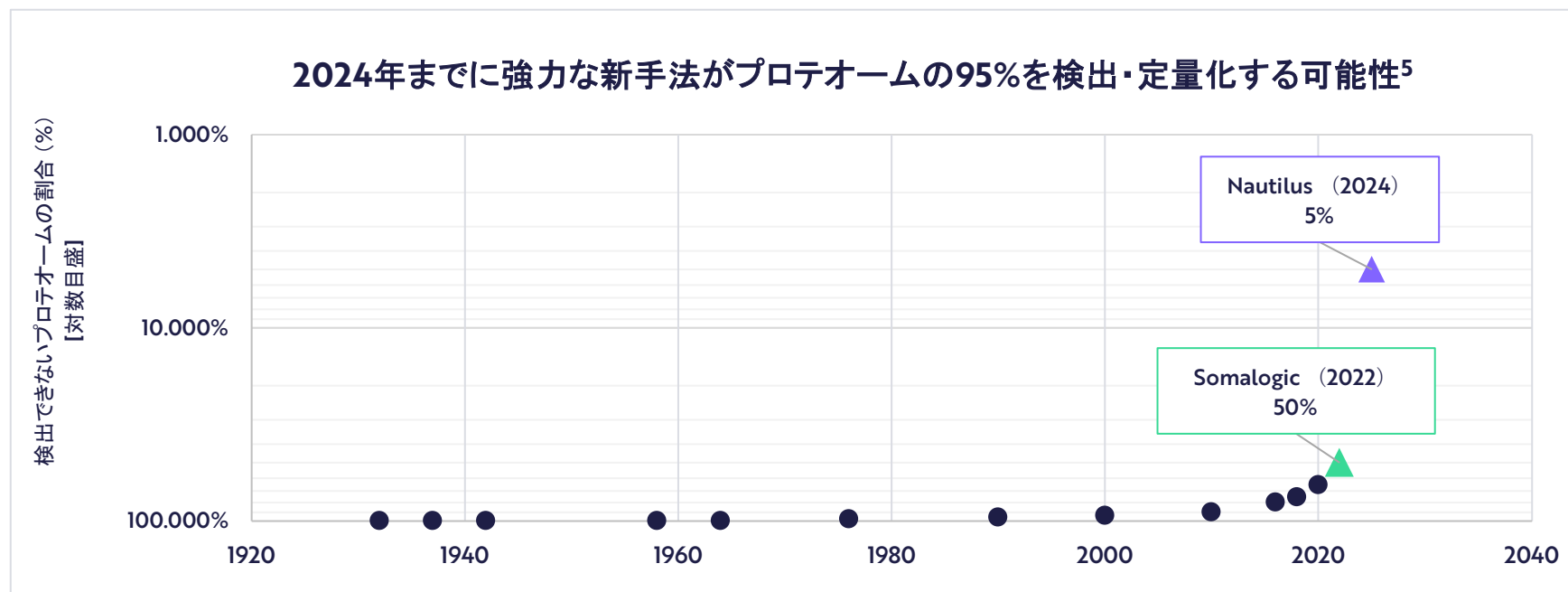
[3] Messner, Christoph B., et al. "Ultra-Fast Proteomics with Scanning Swath." *Nature News*, Nature Publishing Group, 25 Mar. 2021, <https://www.nature.com/articles/s41587-021-00860-4>.

[4] Biosciences, OpenPIP | InterVenn. "OpenPIP: Intervenn Biosciences." *OpenPIP | InterVenn Biosciences*, <https://intervenn.com/enabling-technologies/openpip/>.



## 科学者たちは、これまで発見されなかったタンパク質を検出する方法を創出中

NGSはDNAやRNAデータの解析の並列処理能力を著しく高め、研究実験の大幅な増加を可能にしました<sup>1</sup>。新しい検出技術の一部は、タンパク質の検出と分析を並列的に行なうことを可能にしました<sup>2,3,4</sup>。ARKでは、2022年末までに、科学者が1回の実験でヒトの血漿プロテオームの半分以上\*を定量化できるようになると予測しています<sup>2</sup>。



- 一つのタンパク質が原因の疾患は少なく、複数のタンパク質を検査することが必要とされています<sup>6</sup>
- 多くの遺伝子を含むNGSベースの検査は、タンパク質の「パネル」への道を開くかもしれません
- 多変量解析(ML)とバイオインフォマティクスは、科学者が急増するプロテオミクスデータを理解するのに役立ち、概念実証から臨床への移行を加速させる可能性があります

\*次のスライドで詳しく説明するプロテオフォームではなく、標準構造のヒトタンパク質(1万以上のタンパク質)の半分を指します。

上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] "DNA Sequencing Costs: Data." *Genome.gov*, <https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data>.

[2] "Technology." *SomaLogic*, 15 Nov. 2021, <https://somallogic.com/technology/>.

[3] "Technology." *Nautilus Biotechnology*, <https://www.nautilus.bio/technology/>.

[4] "Olink Explore 3072." *Olink*, 29 Nov. 2021, <https://www.olink.com/products-services/explore/>.

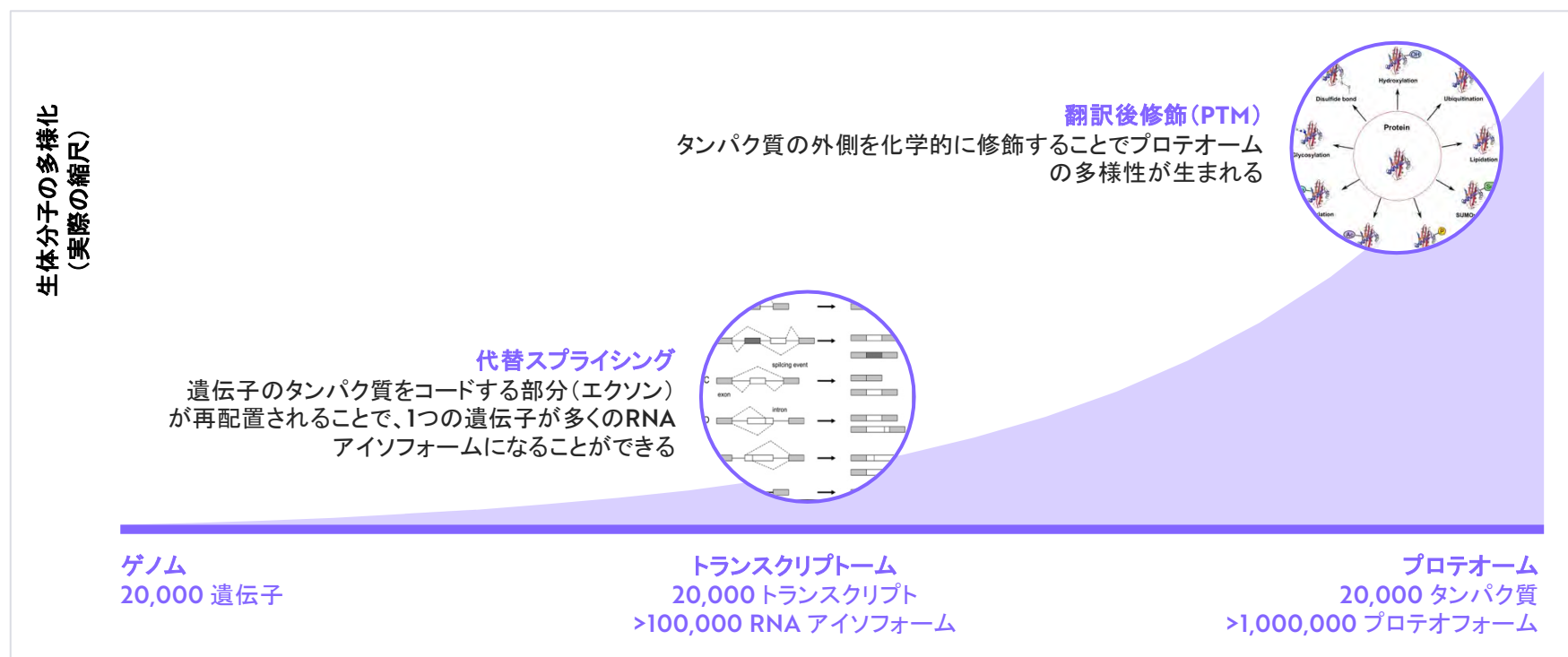
[5] ARK Investment Management LLC, 2021, | the Journal of Molecular and Cellular Proteomics

[6] Anderson, Norman, and N. Leigh Anderson. "The Human Plasma Proteome." *Molecular and Cellular Proteomics*, [https://www.mcponline.org/article/S1535-9476\(20\)34246-8/fulltext](https://www.mcponline.org/article/S1535-9476(20)34246-8/fulltext).



## セントラルドグマの各過程が複雑さを増大させる

ゲノムに含まれる約20,000個の遺伝子が約20,000個の標準的なタンパク質を作り出す一方で、トランスクリプトームやプロテオームには何十万もの異なった微分子が含まれています<sup>1</sup>。例えば、代替スプライシングや翻訳後修飾などの生物学的プロセスは、以下に示したように、多様性と複雑性をもたらします。



- ARKでは、半導体やナノポアを用いたシーケンス技術が、完全長のRNAアイソフォーム、タンパク質の配列変異、および代替スプライシングやPTMなどのプロセスを分析するのに最も適していると考えています<sup>2,3,4</sup>
- 半導体およびナノコアベースのシーケンス技術のスケラビリティと成熟度は、新しい研究結果を臨床に結びつけるのに役立つかもしれません
- **AlphaFold v.2**のようなアルゴリズムは、タンパク質の配列データをプロテオフォーム構造の正確な3D予測に変換することができ、仮説や創薬への道筋を迅速かつ安価に作ることができます<sup>5</sup>

上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] September 12, 2017, et al. "Intro to Proteogenomics." *National Cancer Institute*, <https://www.cancer.gov/about-nci/organization/ccg/blog/2020/intro-proteogenomics-central-dogma>.

[2] Al'Khafaji, Aziz M., et al. "High-Throughput RNA Isoform Sequencing Using Programmable Cdna Concatenation." *BioRxiv*, Cold Spring Harbor Laboratory, 1 Jan. 2021, <https://www.biorxiv.org/content/10.1101/2021.10.01.462818v1>.

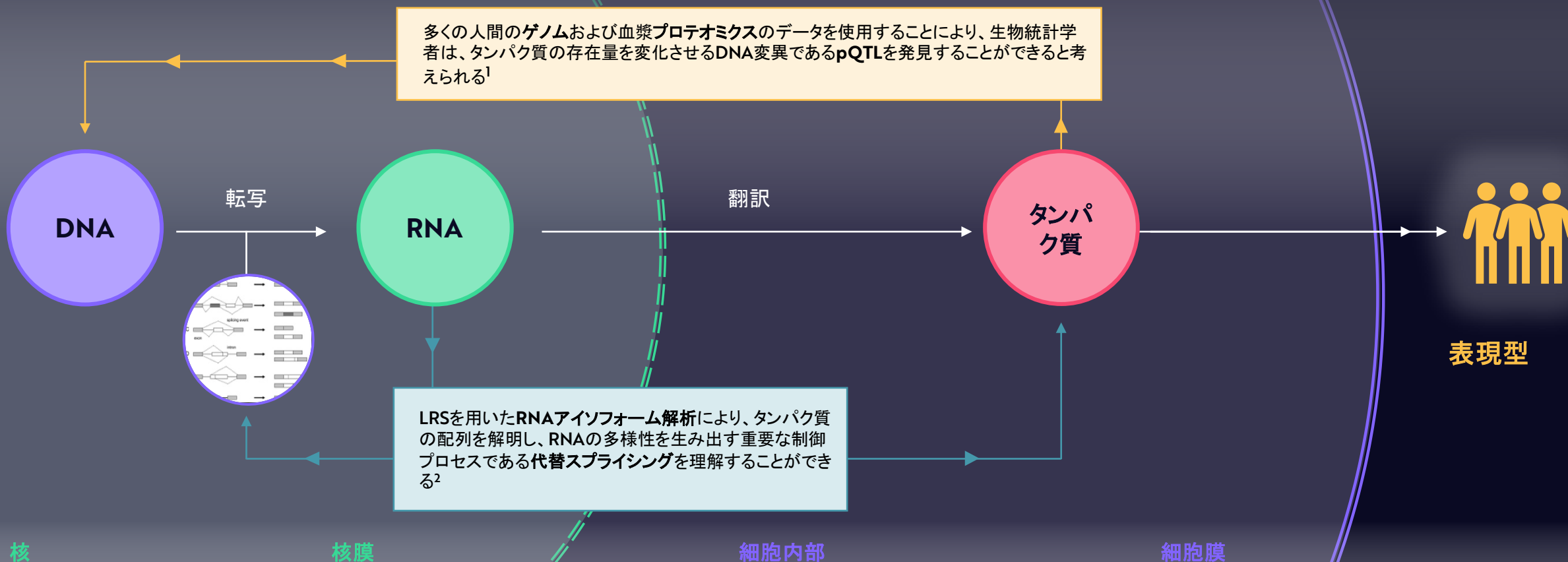
[3] "See beyond Color - Quantum-Si Products and Technology." *Quantum*, 21 June 2021, <https://www.quantum-si.com/products-and-technology/>.

[4] <https://orcid.org/0000-0003-2083-6027>, Winston Timp, et al. "Beyond Mass Spectrometry, the next Step in Proteomics." *Science Advances*, 10 Jan. 2020, <https://www.science.org/doi/10.1126/sciadv.aax8978>.

[5] Jumper, John, et al. "Highly Accurate Protein Structure Prediction with AlphaFold." *Nature News*, Nature Publishing Group, 15 July 2021, <https://www.nature.com/articles/s41586-021-03819-2>.



# 生物学の謎を解き明かす新しいツールと方法



上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] Malone, Cindy S. "Central Dogma of Genetics - CSUN.EDU." [csun.edu, https://www.csun.edu/~cmalone/pdf360/Ch13-1transcription.pdf](https://www.csun.edu/~cmalone/pdf360/Ch13-1transcription.pdf).





# マルチオミクス市場の規模は 1,100億米ドルから3,000億米ドルに拡大する可能性

ARK では、分子生物学の未来は、セントラルドグマの根本部分を統合したマルチオミクス技術にあると考えています。

ARKの研究によると、生命科学分野、基礎研究、トランスレーショナルリサーチ、人々の健康管理、分子診断などを含むマルチオミクスは、がん治療、臓器の診断、人間の健康にインパクトを与え、市場規模は今後5年間は年率22%のペースで増加し、1,100億米ドルから約3,000億米ドルに拡大するとみられます<sup>1</sup>。

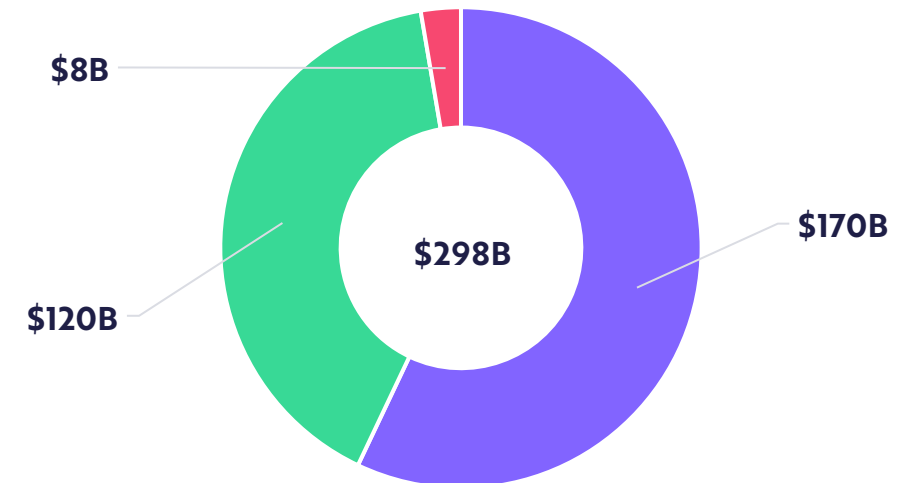
コストの低下に伴い、DNA、RNA、タンパク質を組み合わせた分子診断テストは、既存のテストよりも正確かつ包括的に病気を診断できるようになると考えられます。

ARKでは、ゲノミクス、トランスクリプトミクス、プロテオミクス技術のコストが低下することにより、今後5年間で、**がん、臓器の診断、多くの人間での研究において液体生検が飛躍的に拡大すると予想しています。**

マルチオミクスの市場規模予測

(単位:10億米ドル)

- Oncology
- Organ Health
- Population Health



MRD = Minimal Residual Disease (微小残存病変) の略

上記の予測は、限定的なものであり、その信頼性を保証するものではありません。投資助言を提供するものでも、特定の銘柄の売買や保有を推奨するものでもなく、説明のみを目的としたものです。

[1] ARK Investment Management LLC, 2021 | Surveillance, Epidemiology, and End Results (SEER), SomaLogic (SLGC), and the American Society of Human Genetics (ASHG)