

Annual Review

腎 臟

編集 | 富野康日己 順天堂大学教授
柏原 直樹 川崎医科大学教授
成田 一衛 新潟大学教授

Ann Intern Med
New Engl J Med
Circulation JAMA
Lancet Endocr Rev Science
Ann Neurol
Ann Rev Biochem
Ann Surg Gastroenterology
Cell J Natl Cancer I
Nature Annu Rev Neurosci
Annu Rev Immunol

中外医学社

2013

4. 次世代シーケンサー

新潟大学大学院医歯学総合研究科腎・膠原病内科学分野講師 後藤 眞
 国立遺伝学研究所総合遺伝研究系人類遺伝研究部門 細道一善
 新潟大学大学院医歯学総合研究科腎・膠原病内科学分野教授 成田一衛

key words next generation sequencer, exome analysis, mutation, personal genome

動 向

次世代シーケンサーの登場によりゲノム科学や疾患研究は大きく変貌を遂げている。2007年、J.D.ワトソン博士のゲノム配列が個人ゲノムとしては初めて決定され¹⁾、以後、次々と各人種の個人ゲノムが決定された。ヒトゲノムプロジェクトで約13年間に要したとと比較すると、現在では、次世代シーケンサーによる個人ゲノム配列決定はわずか約10日間で可能であり、費用も著しく

低下している(図1)。いわゆる「1000ドルゲノム」が実現されつつあり、パーソナルゲノム研究が爆発的に普及すると予想される。

A. 次世代シーケンサーの登場と研究の進歩

ヒトゲノム計画に代表されるように、従来の塩基配列決定の中心はサンガー法であった。サンガ

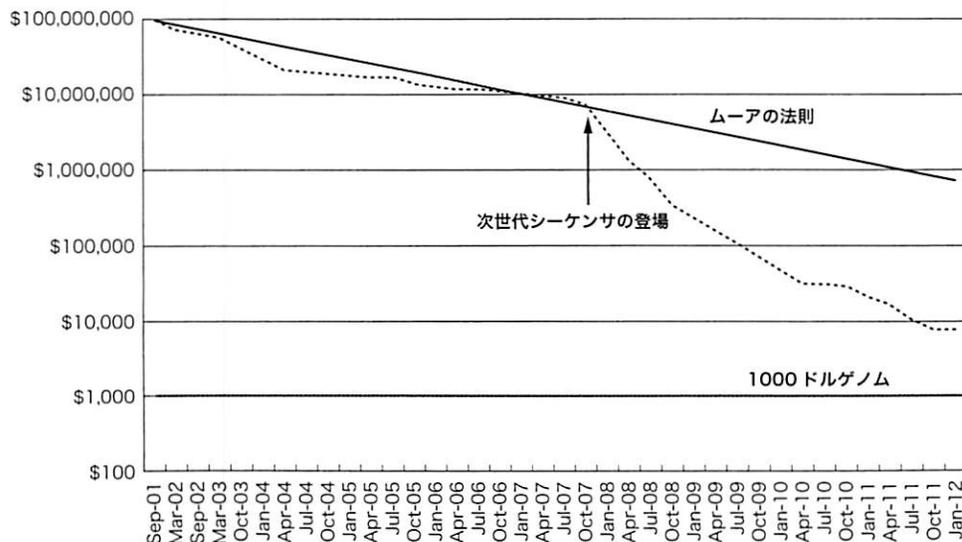


図1 次世代シーケンサーの進歩によるヒトゲノム決定コストの推移

一法の原理は蛍光ラベルされたジデオキシヌクレオチドを用いてDNA伸長反応を止め、電気泳動を行って塩基配列を読み取る。次世代シーケンサーでは、各機種ごとに原理は異なり、1塩基合成シーケンシング法 (HiSeq, イルミナ), パイロシーケンシング法 (454, ロシュ), リガーゼ反応シーケンシング法 (SOLiD, ライフテクノロジー), 蛍光標識および光学的検出方法に頼らず、シーケンス反応を水素イオンで検出するシーケンサー (ionPGM, ライフテクノロジー) などが現在用いられている。また、第三世代といわれる、1分子高速シーケンサー (PacBio RS, Pacific Biosciences) も実用化され、今後もナノポアという、DNA 1分子が通過できる穴を、DNAを通過させ、電流の変化から塩基配列を決定する手法も開発が進んでいる。次世代シーケンサーの特徴はこれらの反応が基板上で超並列で同時に進行するため塩基配列の決定量は飛躍的に向上した。

次世代シーケンサーによる大量かつ高速の塩基配列決定はさまざまな研究分野に影響を与えている。次世代シーケンサーにより可能なアプリケーションとしては、ゲノムアセンブリ、トランスクリプトーム解析 (RNA-Seq など)、リシーケンス解析 (エクソーム解析など)、メチル化・転写因子の解析 (ChIP-Seq など) があげられる。大きく前進した研究領域に癌ゲノム研究と疾患遺伝子研究がある。癌の遺伝子異常は主に体細胞のゲノム配列の異常であり、癌組織と正常組織のそれぞれのゲノム配列を網羅的に比較することで、個々の癌組織での遺伝子の変化をとらえて癌化のメカニズムが解明されることが期待されている。一方、疾患遺伝子研究にもパーソナルゲノム解析が導入され、さまざまな疾患の原因遺伝子や疾患パスウェイが明らかとなっている。

B. 疾患遺伝子研究とエクソーム解析

次世代シーケンサーの導入により、個人の全ゲノム配列を決定して解析するパーソナルゲノム研究が導入されている。従来、単一遺伝子疾患に対してはポジショナルクローニングにより原因遺伝子が同定されてきた。全ゲノム領域を網羅する多型マーカーと疾患の表現型を連鎖解析を用いて原因遺伝子座を決定し、さらに候補遺伝子のシーケンスを行い、原因遺伝子を同定した。しかし、メンデル遺伝形式を示しながら、その家系サイズが小さい場合や、家系数が少ない場合、原因遺伝子を同定することは容易ではなく、原因遺伝子が同定されていない遺伝性疾患が少なからず存在していた²⁾。一方、孤発性疾患に対しては、ゲノムワイド関連解析により疾患遺伝子が同定されてきたが、頻度の高い塩基置換を用いるために、頻度は低いがより効果サイズが大きい疾患遺伝子は同定できなかった。しかしながら、次世代シーケンサーにより大量ゲノム配列の決定、解析が可能となり、特にゲノムの中の全エクソン配列を対照とするエクソーム解析により遺伝性稀少疾患の原因遺伝子を同定したとする報告が多くみられるようになったのである³⁾。

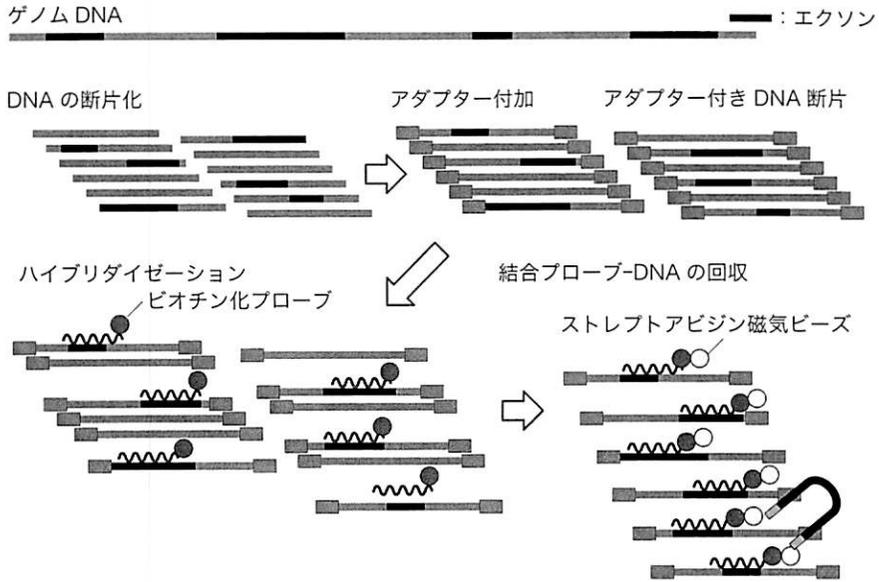
エクソンとはヒトゲノム配列において蛋白質に翻訳される領域であり、長さとしてはヒトゲノムの約1%にあたる30Mb程度であり、エクソン配列に含まれる変異が疾患原因の約85%を説明すると推定されている⁴⁾。エクソンのみを解析することは比較的安価であり、また解析情報量が少ない利点がある。

全エクソンシーケンスは目的ゲノム領域に相補的に設計されたプローブを用いて、ゲノム上の特定領域を濃縮する技術であり、ヒトの全エクソン配列で設計されたビオチン化プローブをDNAライブラリーとハイブリダイズし、結合したプローブ-DNA をストレプトアビジン結合マグネツ

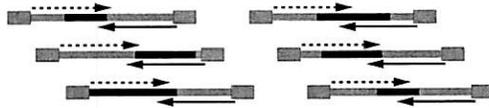
トビーズで回収，結合したDNAを回収することで，シーケンスするためのターゲットDNAを濃縮するというのがその原理である（図2）．データ解析は得られたシーケンスリードをヒトレリファレンス配列 (UCSC hg19) にアライメントし，

シーケンスリードにおけるUCSC hg19との違いを検索することで変異を検出する．これらの変異について遺伝子情報，アミノ酸置換，一塩基置換 (SNP) の公的データベースである dbSNP での登録，1000genome プロジェクトで認められ

1. エクソームシーケンスライブラリーの調整



2. シーケンシング



3. データ解析

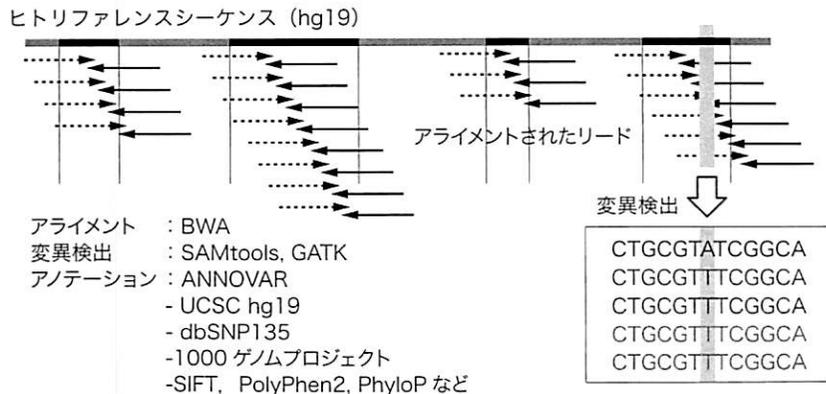


図2 エクソーム解析のワークフロー

ている頻度、アミノ酸置換が表現型に与える影響の予測スコア (SIFT⁵⁾, Polyphen2⁶⁾, LRT, MutationTaster⁷⁾, etc) などの情報を付加する。これらの解析のうちアライメントおよび変異の検出にはオープンソースのプログラムである BWA⁸⁾, Samtools⁹⁾, GATK¹⁰⁾ などが、変異の各種情報の付加には ANNOVAR¹¹⁾ などが用いられる。これら一連の解析により塩基置換や挿入・欠失などの変異リストを作成し、その後の疾患原因遺伝子の絞り込みは、主にフィルタリング法を用いて行われる。エクソーム解析におけるフィルタリングは検出された変異リストに対してどのようなフィルタで絞り込むかが重要な点である。このフィルタリング法では、家系内の罹患者と非罹患者から得られたシーケンズデータから、エクソンにおける非同義置換、スプライス部位および挿入欠失をリスト化し、dbSNPによるありふれた変異の除去、非罹患者との比較などにより候補遺伝子を絞り込む。Ngらは、Miller症候群の罹患者4名について同様の手法を用い、遺伝子変異を同定する手法としてのフィルタリングの有用性を証明した¹²⁾。家系内での遺伝様式が優性遺伝と考えられれば候補となる変異のアリル頻度はおよそ0.5であり、劣性遺伝を想定する場合はおよそ1となる。検出された変異の新規性に関しては、dbSNPへ未登録である、あるいは1000 genome プロジェクトできわめて頻度が低い、または認められないことで絞り込む。また、小規模家系での連鎖解析は、連鎖解析のみの結果だけでは有用な情報とはならないが、フィルタリングにおいては効果的であり、連鎖領域内の候補変異アリル数はかなり限定される。絞り込まれた候補変異の中からさらに原因遺伝子候補を同定するためには、co-segregationを確認する。原因遺伝子変異は罹患者のみに認められ、非罹患者には認められないという、家系内の表現型と変異の有無が一致する必要がある。

全エクソンおよび全ゲノム配列決定による疾患遺伝子研究が加速する一方で、それに伴う問題点も指摘されている。検出される変異が疾患の原因であるか、十分な検討が必要である¹³⁾。疾患遺伝子と同定するには、複数の家系で観察され、かつ遺伝子内に複数の変異が認められる必要がある¹⁴⁾。単一家系のみでの変異については、培養細胞やモデル動物で生物学的な検証が求められる。

C. 次世代シーケンサーを用いた腎臓病研究

次世代シーケンサーを用いたエクソーム解析は、近年、腎臓領域の分野でも応用された例が報告されている。Boydenらは、偽性低アルドステロン症II型 (PHAII) の家系解析にエクソーム解析を応用した¹⁵⁾。PHAIIは、高血圧、高カリウム血症、代謝性アシドーシスを呈し、家族内集積を呈する。過去、PHAIIの家系解析では、WNK kinase が原因遺伝子として同定され、遠位尿細管でNaの持続的再吸収が体液量増加を生じ、高血圧を呈することが明らかとなった¹⁶⁾。しかし、*WNK1* あるいは *WNK4* において遺伝子変異が認められる家系は13%にしか過ぎず、彼らは11家系から *WNK* 変異陰性のPHAII症例をそれぞれ選択し、エクソーム解析により複数の家系でco-segregationを示す28変異を検出した。その中でも *KLHL3* は3家系で5変異が認められ、PHAII 全ての家系における *KLHL3* 変異も、699名のコントロールと比較して有意に多かった。さらに、*KLHL3* と相互作用する *CUL3* にも着目し、*CUL3* のエクソン9のスプライシング異常をきたす変異を、*WNK1*, *WMK4*, *KLHL3* 変異陰性のPHAII家系で見出した。*KLHL3* 変異には優性遺伝形式と劣性遺伝形式が混在しており、従来の遺伝様式を仮定した連鎖解析では原因遺伝子変異を同定す

ることが困難であったことが窺われる。

Sanna-Cherchiらは、ステロイド抵抗性ネフローゼ症候群を呈する家系解析にエクソーム解析を行った¹⁷⁾。近親婚のある常染色体劣性遺伝を示すネフローゼ症候群の1家系に対し、ホモ接合体マッピングとエクソーム解析を行い、候補遺伝子として*MYO1E*と*NEIL1*でのアミノ酸置換を伴う変異を検出した。Myo1eは糸球体上皮細胞の骨格蛋白であり、Myo1eを不活化したマウスではネフローゼ症候群が生じることが示されている。その後、Meleらは複数の小児家族性ネフローゼ症候群の家系に対し、全ゲノム連鎖解析を行い、候補領域について次世代シーケンサーを用いたりシーケンズ解析を行った。上記と同様に*MYO1E*の変異を検出し、小児の家族性巣状糸球体硬化症の原因遺伝子であることを報告した¹⁸⁾。

次世代シーケンサーを用いて多数例の多発性嚢胞腎患者の遺伝子診断が試みられている。常染色体優性多発性嚢胞腎 (ADPKD) は*PKD1*あるいは*PKD2*の遺伝子変異で生じるが、重複偽遺伝子や高度のアレル異質性によりその同定は容易ではない。RossettiらはADPKD患者のDNAから*PKD1*および*PKD2*の各エクソン領域をバーコード化して増幅し、プールの後に、次世代シーケンサーによりゲノム配列を決定した¹⁹⁾。183名のADPKD患者の中で115名 (63%) に病的変異が認められ、その約10%は従来のサンガー法では検出が困難であった変異 (allele dropoutやgene conversionなど) であり、今後のADPKDの遺伝子診断のモデル解析になり得ると報告している。

ネフロン癆は腎髄質に嚢胞形成を認める遺伝性疾患であり、繊毛機能異常を伴い、腎外徴候も認められる。遺伝的異質性 (locus heterogeneity) が高度であり、18遺伝子に原因変異が存在すると報告されており、変異解析は非常に時間と労力を要する作業である。Ottoらは、24名の患者の

DNAをプールし、18遺伝子の376エクソンを全て含む領域をPCRで増幅し、次世代シーケンサーで変異を解析した²⁰⁾。その結果、120名の患者の中で30名 (25%) に54の病的変異を検出することが可能であり、大量リシーケンズ解析の有用性を報告した。

トランスクリプトーム解析にも次世代シーケンサーを用いた研究が報告されている。Brennanらは、腎線維化のモデルとしてHK-2培養細胞にTGF- β 1を加え、RNA-Seqを用いた網羅的な遺伝子発現を解析した²¹⁾。NF κ Bなどの比較的少数の転写因子が線維化を進める遺伝子発現を制御していることが明らかとなり、マイクロアレイ解析では同定されていなかった低発現の遺伝子群が線維化の進行に関与していることも明らかにされた。

D. パーソナルゲノム研究と倫理

次世代シーケンサーによる技術革新により、個人のゲノム配列は短期間で決定できる時代となり、いわゆる個人化医療として大量のゲノム配列情報が臨床の現場へ応用される可能性が高まってきた。応用される分野としては、疾患の診断、薬剤に対する応答性、副作用の起こり易さ、などが考えられる。

個人の全ゲノム解析を行う上で、重要な点は倫理的な配慮である。疾患遺伝子を同定するためには罹患者だけでなく、家系内あるいは一般集団でのゲノム配列情報が必要になる。ゲノム配列情報は最も重要な個人情報であり、ゲノム研究を行う際にはDNA提供者へ個人情報が十分に保護されることを含めた説明がなされなければならない²²⁾。パーソナルゲノムの進展は、医療一般および社会全体が予想するよりも早く進む可能性があり、十分な基盤整備が望まれる。

文献

- 1) Wheeler DA, Srinivasan M, Egholm M, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature*. 2008; 452: 872-6.
- 2) Gilissen C, Hoischen A, Brunner HG, et al. Unlocking Mendelian disease using exome sequencing. *Genome Biol*. 2011; 12: 228.
- 3) Yamaguchi T, Hosomichi K, Narita A, et al. Exome resequencing combined with linkage analysis identifies novel PTH1R variants in primary failure of tooth eruption in Japanese. *J Bone Miner Res*. 2011; 26: 1655-61.
- 4) Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc Natl Acad Sci USA*. 2009; 106: 19096-101.
- 5) Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009; 4: 1073-81.
- 6) Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7: 248-9.
- 7) Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. *Hum Mutat*. 2011; 32: 894-9.
- 8) Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25: 1754-60.
- 9) Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009; 25: 2078-9.
- 10) McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*. 2010; 20: 1297-303.
- 11) Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010; 38: e164.
- 12) Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet*. 2009; 42: 30-5.
- 13) Lyon GJ, Wang K. Identifying disease mutations in genomic medicine settings: current challenges and how to accelerate progress. *Genome Med*. 2012; 4: 58.
- 14) Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011; 12: 745-55.
- 15) Boyden LM, Choi M, Choate KA, et al. Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature*. 2012; 482: 98-102.
- 16) Wilson FH, Disse-Nicodème S, Choate KA, et al. Human hypertension caused by mutations in WNK kinases. *Science*. 2001; 293: 1107-12.
- 17) Sanna-Cherchi S, Burgess KE, Nees SN, et al. Exome sequencing identified MYO1E and NEIL1 as candidate genes for human autosomal recessive steroid-resistant nephrotic syndrome. *Kidney Int*. 2011; 80: 389-96.
- 18) Mele C, Iatropoulos P, Donadelli R, et al. MYO1E mutations and childhood familial focal segmental glomerulosclerosis. *N Engl J Med*. 2011; 365: 295-306.
- 19) Rossetti S, Hopp K, Sikkink RA, et al. Identification of gene mutations in autosomal dominant polycystic kidney disease through targeted resequencing. *J Am Soc Nephrol*. 2012; 23: 915-33.
- 20) Otto EA, Ramaswami G, Janssen S, et al. Mutation analysis of 18 nephronophthisis associated ciliopathy disease genes using a DNA pooling and next generation sequencing strategy. *J Med Genet*. 2011; 48: 105-16.
- 21) Brennan EP, Morine MJ, Walsh DW, et al. Next-generation sequencing identifies TGF- β 1-associated gene expression profiles in renal epithelial cells reiterated in human diabetic nephropathy. *Biochim Biophys Acta*. 2012; 1822: 589-99.
- 22) Kato K. [Ethical issues of personal genome research]. *Rinsho Shinkeigaku*. 2011; 51: 975.

p 351

Annual Review ^{じんぞう}腎臓 2013 ©

発行 2013年 1月25日 初版 1刷

編集者 富野 康日己
柏原 直樹
成田 一衛

発行者 株式会社 中外医学社
代表取締役 青木 滋

〒162-0805 東京都新宿区矢来町62
電 話 03-3268-2701 (代)
振替口座 00190-1-98814 番

印刷・製本 / 東京リスマチック (株) <KS・YT>
ISBN978-4-498-12483-7 Printed in Japan

JCOPY <(社) 出版者著作権管理機構 委託出版物>

本書の無断複写は著作権法上での例外を除き禁じられています。
複写される場合は、そのつど事前に、(社)出版者著作権管理機構
(電話 03-3513-6969, FAX 03-3513-6979, e-mail: info@jcopy.
or.jp) の許諾を得てください。