



がん全ゲノム解析への展開

間野 博行[†]

IRYO Vol. 74 No. 11/12 (508-510) 2020

【キーワード】 全ゲノムシーケンシング, 構造異常, クリニカルシーケンシング

要旨

薬剤が紐付いている遺伝子をまとめて配列解析する「がん遺伝子パネル検査」が2019年6月に保険償還され、日本でもがんゲノム医療が開始された。がん遺伝子パネル検査は、現在ある薬剤を患者に届ける最も効率のよい方法であるが、原因が不明で有効な薬剤が存在しないがん種が今なお数多く存在する。そのようながんの原因を解明するためには、これまであまり解析がなされてこなかったがんの全ゲノム解析が必要であり、大規模に行う際の技術的ハードルも下がってきた。将来のクリニカルシーケンスとしての全ゲノム解析も視野にはいつてきたといえる。

はじめに

がん研究が進むことにより、それぞれのがんにおける本質的な発がん原因遺伝子が明らかになり、その遺伝子産物の機能を抑制する分子標的治療薬が誕生した。さらに大規模にがんゲノム解析を行うプロジェクトが米国(The Cancer Genome Atlas, <https://portal.gdc.cancer.gov>) および国際共同研究(International Cancer Genome Consortium : ICGC, <https://icgc.org>) で行われ、これまで見逃されてき

た数多くのがん関連遺伝子変異が同定された。たとえばICGCでは25,000例のがん検体とペア正常検体とで全エクソンシーケンス解析を行い、がんにおけるタンパク配列の異常を網羅的に探索した。その結果、各がん種における体細胞変異遺伝子のカタログ化が行われ、がんの解明が大きく進んだといえる。

同定された発がん原因遺伝子変異とそれをターゲットとした分子標的薬剤、またそれぞれの遺伝子変異を検出するコンパニオン診断薬が多数出現したことにより、がんを診断する際に「対応する薬剤がある遺伝子変異は一度に検出する」検査としてがん遺伝子パネル検査が出現した。これを用いて、患者毎に最適の治療法を選択する医療が「がんゲノム医療」といえる。

がん遺伝子パネル検査によるゲノム医療は、既知の薬剤関連遺伝子変異を効率よく検出するには最適の手法といえるが、分子標的薬剤が現時点では存在しないがん種に対しては無力である。たとえば、最も分子標的薬剤が豊富な肺腺がんであっても、約1/4の症例については有効な分子標的薬が存在しない(図1)。がん全体でいうと、何らかの標的薬剤の対象となるがんの割合は3割程度に満たない¹⁾。そこで、これまで全エクソンシーケンス解析で見逃されてきたゲノム領域も調べる全ゲノムシーケンス(whole genome sequencing : WGS)が注目されている。

国立がん研究センター 研究所 [†]医師

著者連絡先：間野博行 国立がん研究センター 研究所長 〒104-0045 東京都中央区築地5-1-1

e-mail : hmano@ncc.go.jp

(2020年6月15日受付, 2020年7月10日受理)

Application of Whole Genome Sequencing for Cancer

Hiroyuki Mano, National Cancer Center

(Received Jun 15, 2020, Accepted Jul. 10, 2020)

Key Words : whole genome sequencing, structural variation, clinical sequencing

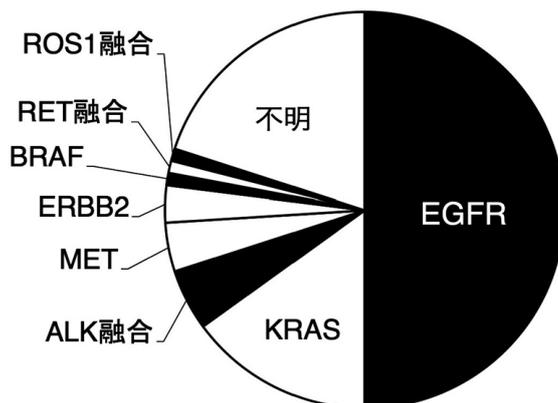


図1 日本人肺腺がんにおけるがん遺伝子
肺腺がんの原因となるがん遺伝子の日本人における頻度を表す。黒色領域の遺伝子は対応する分子標的薬剤が日本で承認済み。

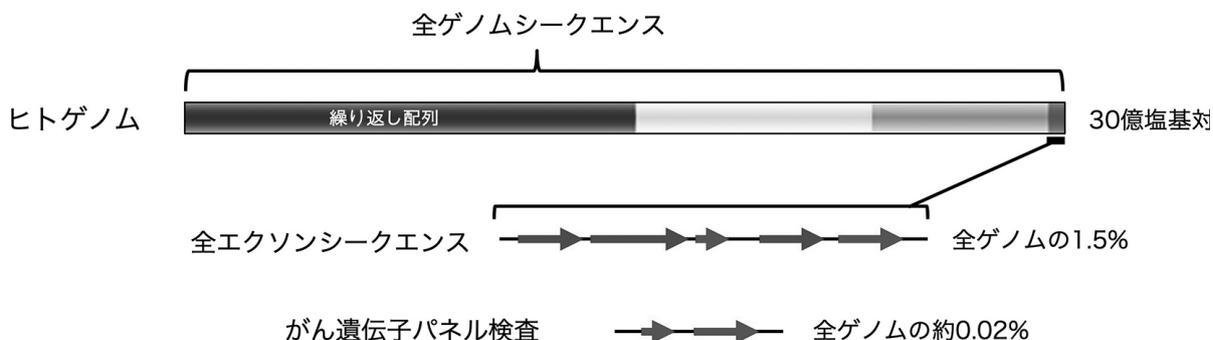


図2 ヒトゲノムの構造

ヒトゲノムはハプロイドあたり約30億塩基対の長さであり、約半分がSINE/LINEなどの繰り返し配列からなる。遺伝子領域は全体の約20%であるが、そのほとんどはイントロンであり、アミノ酸配列をコードするエクソン領域は全体のわずか1.5%程度である。

がん全ゲノム解析

ヒト細胞は46本の染色体を含んでおり、ゲノムDNAの長さはハプロイドあたり約30億塩基対に及ぶ(図2)。ゲノム全体の約半分は繰り返し配列であり、タンパクをコードするエクソン領域はわずか1.5%にすぎない。全エクソンシーケンスはこの領域のみを選択的に解析することで、アミノ酸配列異常を効率よく検出しているのである。また、エクソン領域にある約2万個の遺伝子の内、薬剤が紐付いているがん関連遺伝子のみを解析する手法ががん遺伝子パネル検査である。

WGS解析をすることで、エクソン以外の領域の体細胞変異が発がんに寄与することが明らかになると予想される。実際すでに、エクソン以外の領域で

の発がん機構が存在することが明らかになってきた。Killelaらは、*TERT*遺伝子(テロメラーゼをコードする)のプロモーター領域に一塩基置換が生じることで同遺伝子の過剰発現がもたらされることを発見している²⁾。驚くべきことにこの変異の割合は、膠芽腫、膀胱がんでは過半数に及び、きわめて高頻度ながん遺伝子変異となっている。さらに、*TERT*遺伝子に関しては、上流領域にAAV2ウイルスゲノムが挿入され、テロメラーゼの高発現が生じていることが肝細胞がんにおいて報告されている³⁾。またわれわれは、思春期および若年成人発症の急性リンパ性白血病において、*DUX4*という繰り返し配列の1~数コピーが免疫グロブリン遺伝子座に転座して、過剰発現され白血病を誘導することを見いだした⁴⁾。この世代の急性リンパ性白血病では最も高頻度のがん遺伝子となっており、エクソン解析だけで

はみつからなかった情報である。さらに片岡らは成人T細胞白血病の約3割においてPD-L1遺伝子(CD274)の3'末領域に染色体構造異常が生じてPD-L1の高発現をきたし、白血病細胞が宿主の免疫監視機構から逃れていることを発見した⁵⁾。

さらに多くのがん原因変異がWGSによって明らかになると予想されるが、WGSを大規模に行うにはさまざまな制約もある。米国の大規模全エクソンシーケンシングプロジェクトTCGAでは数百億円の予算が投下されたが、エクソン領域の70倍近いヒト全ゲノムシーケンシングを多数の検体で行うのは国として大きな予算が必要である。またシーケンシング費用だけでなく、それを解析するコンピューターリソース、解析結果を共有するクラウドシステムなど多くの面でインフラ作りが必要である。近年は、旧来のコンピュータチップ(central processing unit: CPU)に替わって、ゲノム解析のためのGPU(graphics processing unit)パイプラインあるいはFPGA(field programmable gate array)などが開発され、解析スピードが飛躍的に増大しているため、以前よりは大規模解析のハードルは下がったといえる⁶⁾⁷⁾。

クリニカルシーケンシングへの展開

現在は、がん遺伝子パネル解析がクリニカルシーケンシング(臨床検査としてのシーケンシング)に用いられており、最も費用対効果の高いゲノム検査法といえる。解析結果が薬剤選択の根拠になるため、できるだけシーケンシングエラーを下げる必要がある。一般のがん遺伝子パネル検査では×400から×500のシーケンシング深度(各塩基を平均400-500回読む)で解析している。したがってシーケンシング費用も相対的に高価となり、薬剤の紐付いているがん関連遺伝子だけを選んでシーケンシングするがん遺伝子パネル検査が効率がよい。さらにがんのクリニカルシーケンシングを行う検体としてはホルマリン固定標本(formalin-fixed paraffin-embedded tissue: FFPE)がほとんどであり、FFPEにおいてはゲノムDNAが300塩基対以下に断片化されている⁸⁾。したがってWGSの大きなアドバンテージである染色体の構造異常の検出力が低下することが多い。

しかし、今後がんのWGSが大規模に行われることにより、新たな発がん原因が次々と同定されると期待され、現在はWGSの費用は高価であるが低下していくと予想される。上述のように計算リソース

の急速な進歩もあるため、WGSが最終的なクリニカルシーケンシングとなる可能性も高い。そのような状況にあって重要性が高くなるのは、各遺伝子変異・ゲノム構造異常に対して適切なアノテーション(各変異の臨床的意義づけ: 病原性情報や対応する臨床試験などを記載する)を付与するがんゲノム医療用の知識データベースの質であろう。

著者の利益相反: 本論文発表内容に関連して申告なし。

[文献]

- 1) Bailey MH, Tokheim C, Porta-Pardo E et al. Comprehensive characterization of cancer driver genes and mutations. *Cell* 2018; **173**: 371-85 e18.
- 2) Killela PJ, Reitman ZJ, Jiao Y et al. TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc Natl Acad Sci US A* 2013; **110**: 6021-6.
- 3) Nault JC, Datta S, Imbeaud S et al. Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat Genet* 2015; **47**: 1187-93.
- 4) Yasuda T, Tsuzuki S, Kawazu M et al. Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. *Nat Genet* 2016; **48**: 569-74.
- 5) Kataoka K, Shiraishi Y, Takeda Y et al. Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. *Nature* 2016; **534**: 402-6.
- 6) Franke KR, Crowgey EL. Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for Genome Analysis Toolkit algorithms. *Genomics Inform* 2020; **18**: e10.
- 7) Ji J, Shen L, Bootwalla M et al. A semiautomated whole-exome sequencing workflow leads to increased diagnostic yield and identification of novel candidate variants. *Cold Spring Harb Mol Case Stud* 2019; **5**: a003756.
- 8) Dietrich D, Uhl B, Sailer V et al. Improved PCR performance using template DNA from formalin-fixed and paraffin-embedded tissues by overcoming PCR inhibition. *PLoS One* 2013; **8**: e77771.