

ミニレビュー

乳腺科学分野ハイライト

がんゲノムの多様性解析～ mutational signature からがんのバイオロジーを探る～

Diversity analysis of cancer genome ; Exploring cancer biology from mutational signature

東京医科大学乳腺科学分野：浅岡真理子

The Department of Breast Surgery,

Tokyo Medical University

Mariko ASAOKA

がんゲノムとバイオロジー

がんは遺伝子の疾患であり、正常の遺伝子 (DNA) に傷 (突然変異) が蓄積した結果としてがんが生じるとされており、細胞内で生じた体細胞突然変異は、細胞分裂のときの DNA 複製の誤り、外因性あるいは内因性発がん物質の暴露、DNA の酵素的修飾、DNA 修復系の異常といったものに起因する。一部のがんにおいては、体細胞突然変異の主要因が特定の発がん物質への暴露や DNA 修復系の異常であり、また特徴的な遺伝子変異を伴うことが解明された。この特徴的な遺伝子変異を標的とした分子標的薬が開発され、実臨床の場でも多数使用されている。例えば、EGFR 遺伝子変異を持つ肺癌には EGFR 阻害薬 (Gefitinib)、Her2 遺伝子変異を持つ乳癌には Her2 タンパク質に対する抗体薬 (Trastuzumab) の分子標的治療があり、それぞれ好成績を収めている。このように一部のがんでは研究が重ねられ、発がん要因から遺伝子変異の解明、さらには治療薬の開発にまで及んでいるが、多くのがんでは詳細はいまだ不明である。どの遺伝子で、どのような過程で、どのような遺伝子変異が生じているのかを解明し、遺伝子変異スペクトラムをカタログ化することは、がんに対する新規治療のみならず、早期発見、早期診断、予防法を開発するための重要な鍵となると考えられている。

がんゲノム解読による新たながんの理解

がんにおける遺伝子変異の研究は、1953 年、ジェームズ・ワトソンとフランシス・クリックにより分子模型を構築する手法を用い DNA の二重らせん構造が提唱されたことから始まった。細胞分裂の過程で、DNA は複製され新規細胞内にも同一の遺伝情報が受け継がれることから、遺伝子塩基配列が全 DNA 情報を担っていることが説明できるようになった¹⁾。1969 年、EM Witkin らは、皮膚癌細胞の DNA において分子学的に C > T (シトシンがチミンに単塩基変換される) または CC > TT の連続した 2 塩基置換が生じていることを発見した²⁾³⁾。これが外的因子 (環境因子) である紫外線と DNA 損傷 (突然変異) の関連性を示唆する史上初の報告となり、がんの研究分野にパラダイムシフトを引き起こした。その後も紫外線のような外的因子と遺伝子変異についての関連性が数多く報告され、1977 年 Sanger らによる DNA シークエンス法 (Sanger 法)⁴⁾が開発されたことを発端に、がんのゲノム解析は急速に発展した。一度細胞内で生じた体細胞突然変異は、次世代の細胞へと引き継がれていくことから、“Carcinogens leave fingerprints” (発がん性物質による DNA 損傷はがん細胞遺伝子内に保持される)⁵⁾ という概念が定着した。

1980 年から 1990 年代のがんゲノム研究は Sanger 法による分析が主流であり、実際にタンパクを指定する遺伝子領域 (エクソン領域) を中心とした解析が行われた。代表的ながん抑制遺伝子である TP53 遺伝子の変異は、すべてのがん腫において最も高頻度に見られる体細胞遺伝子変異であることから、特に重点的に解析された。その結果、同一の遺伝子上であっても各がん腫によって変異スペクトラムの違いがあることが証明された⁶⁾。当時は、ターゲットとする一つあるいは数個の遺伝子についての研究がなされていたが、“がんという難解な遺伝子病を理解するには、膨大な数の遺伝子の中から個々の遺伝

子についてそれぞれ研究するのではなく、ヒトゲノム全体を把握することが不可欠である。そのためにはヒトゲノム配列の全てを決定するのが何よりも早道だ。”ノーベル賞受賞者 Renato Dulbecco の発言よりヒトゲノムプロジェクト（ヒトゲノムの全 30 億塩基配列を決定し遺伝情報を解析する計画）が国際ヒトゲノム配列コンソーシアムによって立ち上げられた。DNA 解読技術の飛躍的な進歩に伴い、2001 年最初のヒトゲノム配列が決定され⁷⁾⁸⁾、それ以来、高速次世代シーケンシング技術と情報分析ツールの急速な進展によりがん研究分野は目まぐるしく発展した。

2012 年、Govindan R らにより、喫煙者由来と非喫煙者由来の肺癌腫瘍を比較すると、喫煙者由来の腫瘍内では体細胞突然変異の頻度が 10 倍以上に上り、中でも C>A 塩基置換が圧倒的多数を占めることが報告された⁹⁻¹¹⁾。紫外線を原因とした皮膚癌の塩基変置換はピリミジン領域の C>T または CC>TT 塩基置換が生じる一方で、末端黒子型黒色腫では CpG 領域（シトシンの次にグアニンが現れるタイプの 2 塩基配列）における C>T 塩基置換が主体となることを発見した¹²⁻¹⁴⁾。Matthew J Ellis らは、未治療の ER 陽性 HER2 陰性乳癌の生検検体を用いて、細胞周期や薬剤感受性、がんの悪性度に関わる遺伝子変異を検索した。MAP3K1 遺伝子変異は ER 陽性 HER2 陰性乳癌の中でも特に比較的予後が良好とされている Luminal A タイプや組織学的低 Grade、細胞増殖能の低いタイプに高頻度に認められた。GATA3 遺伝子変異を有する乳癌では、アロマターゼ阻害薬治療によりがん細胞の細胞周期が抑制されることが証明された。臨床学的には同一に分類されていた ER 陽性 HER2 陰性乳癌でも、特異的な体細胞変異パターンによって、全く異なる生物学的特徴を持つことが示された¹⁵⁾。このように個々のサンプルから遺伝子情報を抽出し、臨床学的データと照らし合わせることで、目的とした遺伝子にどのような変異が生じたのかを観察することが可能であった。しかし、体細胞変異は受精卵の最初の分割以降、様々な内因的外因的刺激を受けて遺伝子変異が蓄積されるため、同一腫瘍を構成するがん細胞にも不均一性 (Intratumoral heterogeneity) が生じる。さらには腫瘍増大や腫瘍微小環境からの影響、治療介入などにより選択圧 (Selective pressure) が加わることで、腫瘍組織内における多様性に变化加わる。

こうした要因から、ひとつの腫瘍のなかには異なるゲノムを有する複数のクローンが存在し、様々なプロセスにより生じた腫瘍内の全遺伝子変異を解明することは従来困難であった。

がんゲノム研究の飛躍的展開

2013 年、国際がんゲノムコンソーシアム (International Cancer genome Consortium: ICGC) 内の大規模国際共同研究によって、30 癌腫、7,000 例以上の全ゲノムまたはエクソン領域の DNA シークエンスを行い、約 500 万個の体細胞変異を抽出し網羅的解析が行われた¹⁶⁾。がん腫内に認められる 6 種類の単一塩基置換 (C・G>A・T, C・G>G・C, C・G>T・A, T・A>A・T, T・A>C・G, T・A>G・C) を特定し、さらに周囲の連続した塩基配列の違いも含めた 96 種の塩基置換パターンに分類した。非負値行列因子分解 (NMF: Nonnegative Matrix Factorization) 解析手法を用いて、各がん腫内においてどのような体細胞突然変異パターンの組み合わせ (mutational signature) がどの程度貢献しているのかについて検討された。本研究により 21 種類の mutational signature が同定され、個々のがん腫において最も優勢な Signature をもとに 30 種のがん腫と各 mutational signature の関連性が示された (Figure 1)。多くのがん種では 2 つ以上のパターンの混在が見られ、最大で肝臓癌、胃癌、子宮癌において 6 種類のパターンの混在が認められた。同定された遺伝子変異スペクトラムについては The Catalogue of Somatic Mutations In Cancer (COSMIC) の mutational signature として、Web 公開されている¹⁷⁾。

単一塩基置換に注目すると、Signature1A/1B、6、7、11、15、19 では C>T 置換、Signature4、8、18 では C>A 置換、Signature5、12、16、21 では T>C 置換、Signature9、17 は T>G 変異がそれぞれ優位に認められた (Figure 2)。

Signature1 は全がん腫において最も高頻度に見られた変異スペクトラムであり、生理現象として起こる 5-メチルシトシンの脱アミノ化によって引き起こされることから、がん診断時の年齢に関与すると考えられている。

Signature2 は TpCpN (N はどの塩基でも可) 領域における C>T または G 変異を生じる Signature であり、30 癌腫中 16 種で確認され特に子宮頸癌と腎癌で多くみられた。Signature13 は大多数が Signature2

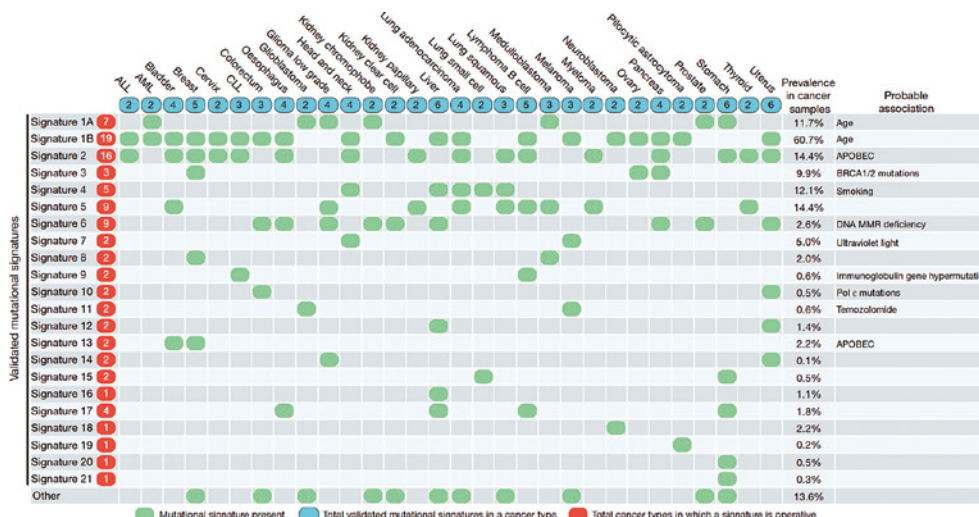


Figure 1 ヒトにおける各がん種と COSMIC mutational signature (Version2) との関連性
Prevalence in cancer samples (%) は、mutational signature に寄与した検体の割合を示す。
<https://cancer.sanger.ac.uk/cosmic/signatures> より出典

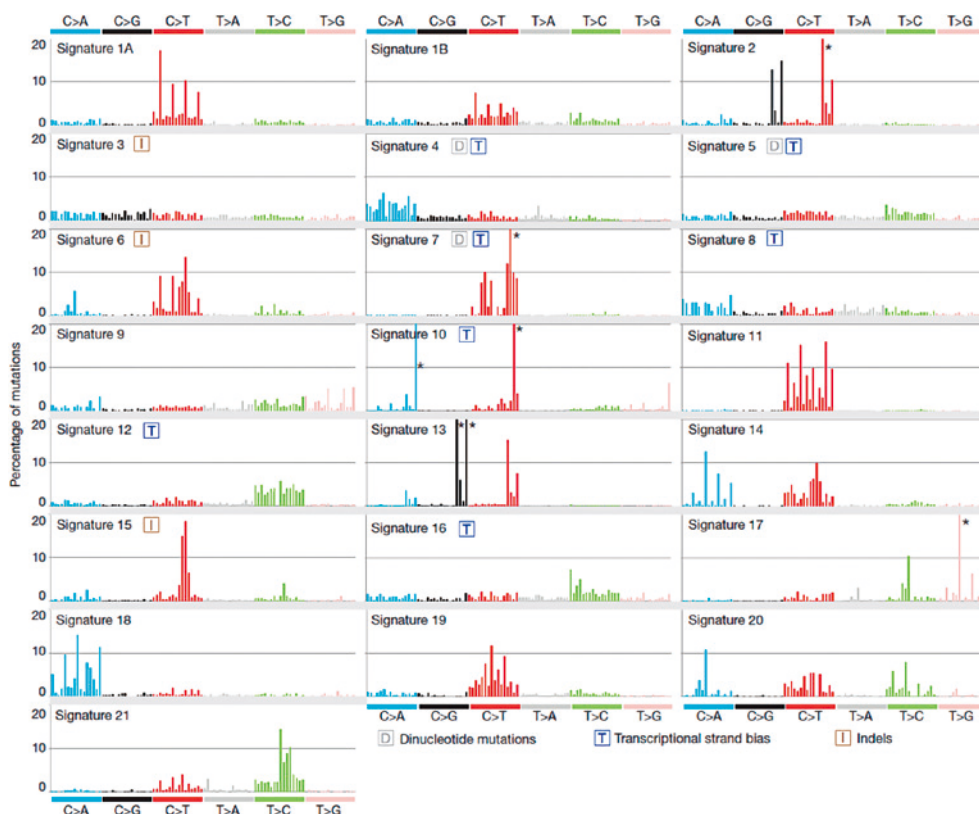


Figure 2 各 mutational signature の変異スペクトラム
横軸は mutational signature 中における単塩基置換遺伝子変異の種類を示す。縦軸は全遺伝子変異に対する各単塩基置換遺伝子変異の割合 (%)、*はその単塩基置換遺伝子変異の割合が 20% 以上であることを示す。
<https://cancer.sanger.ac.uk/cosmic/signatures> より出典

と同一腫瘍内に混在しており、また変異スペクトラムの相同性から、APOBEC (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) という DNA 複製や DNA 修復時に C>U (ウラシル) の変異導入

機能を持つ酵素群による反応によって生じる変異であると推測される。

Signature3 は、乳癌、卵巣癌及び膀胱癌の癌腫内に認められ、BRCA1 または BRCA2 遺伝子変異に関

するものと考えられている。DNA 相同組み換えの際の DNA2 重鎖切断修復時に生じた変異であり、3 塩基対以上の欠損又は挿入が DNA 切断個所に重複して認められた。また、Signature7 は主に皮膚癌で特異的であることから紫外線との関連が、Signature4 は肺癌や頭頸部癌で認めることから喫煙による有害物質との関連がそれぞれ推測され、既知の発がん性物質に伴う遺伝子変異スペクトラムが解明された。各 mutational signature の変異スペクトラムは、COSMIC で公開されている¹⁷⁾。

2020 年、国際がんゲノムコンソーシアム (ICGC ; International Cancer Genome Consortium) およびがんゲノムアトラス (TCGA ; The Cancer Genome Atlas) のがん種横断的全ゲノム解析 (PCAWG ; Pan-Cancer Analysis of Whole Genomes) コンソーシアムにより、これまでで最大の 38 種類のがん 2,658 症例の統合解析結果が発表され、約 4,400 万個の一塩基変異、約 240 万個の短い欠失・挿入、約 29 万個の構造異常、約 8,000 個のミトコンドリアゲノム変異が同定された。また、がんゲノムにはエクソン領域およびイントロン領域を合わせて平均 4~5 個のドライバー変異が含まれていたが、希少癌など一部においてはドライバー変異の特定には至らなかった¹⁸⁾。また新規の mutational signature として、49 種の単塩基置換 (SBS)、11 種の二塩基置換 (DBS)、4 種のクラスター塩基置換、17 種の塩基欠損・挿入 (Indel) が同定され、各癌腫との関連性が示された。

さらには、複数のがん腫において、スプライシング・発現レベル・融合遺伝子・プロモーター活性といった点において体細胞変異が転写に様々な影響を及ぼしうることなど、がんゲノムの多様な全体像が明らかとなった。さらに、今回の解析は、がんゲノムを特徴付ける変異過程の性質とタイミングを明らかにする上で役立つと考えられ、がんの早期発見の機会になることも示唆されている¹⁹⁾²⁰⁾。

がんゲノム研究と今後の展望

このように、一部の mutational signature ではその要因や関連するがん腫が解明された一方で、遺伝子変異スペクトラムは把握されても臨床学的意義にまだ達していないものも多く見受けられる。がんゲノムのビックデータと臨床疫学情報を組み合わせることで、疾患の原因究明に関する未知の発見から発癌経路の解明研究、あるいは個別化医療において

有用なゲノム異常の同定、薬剤治療効果判定に代用可能なバイオマーカー検索や予後因子など生物学的メカニズムの解明に寄与することが期待される²¹⁻²⁴⁾。

文 献

- 1) Watson JD, Crick FH: Molecular structure of nucleic acids ; a structure for deoxyribose nucleic acid. *Nature* **171** : 737-738, 1953. doi : 10.1038/171737a0.
- 2) Witkin EM : Ultraviolet-induced mutation and DNA repair. *Annu Rev Microbiol* **23** : 487-514, 1969. doi : 10.1146/annurev.mi.23.100169.002415.
- 3) Howard BD, Tessman I: IDENTIFICATION OF THE ALTERED BASES IN MUTATED SINGLE-STRANDED DNA. II. IN VIVO MUTAGENESIS BY 5-BROMODEOXYURIDINE AND 2-AMINOPURINE. *J Mol Biol* **9** : 364-371, 1964. doi : 10.1016/s0022-2836(64)80213-8.
- 4) Sanger F, Nicklen S, et al. : DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74** : 5463-5467, 1977. doi : 10.1073/pnas.74.12.5463.
- 5) Vogelstein B, Kinzler KW: Carcinogens leave fingerprints. *Nature* **355** : 209-210, 1992. doi : 10.1038/355209a0.
- 6) Hollstein M, Hergenhahn M, et al. : New approaches to understanding p53 gene tumor mutation spectra. *Mutat Res* **431** : 199-209, 1999. doi : 10.1016/s0027-5107(99)00162-1.
- 7) Venter JC, Adams MD, et al. : The sequence of the human genome. *Science* **291** : 1304-1351, 2001. doi : 10.1126/science.1058040.
- 8) Lander ES, Linton LM, et al. : Initial sequencing and analysis of the human genome. *Nature* **409** : 860-921, 2001. doi : 10.1038/35057062.
- 9) Govindan R, Ding L, et al. : Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell* **150** : 1121-1134, 2012. doi : 10.1016/j.cell.2012.08.024.
- 10) Imielinski M, Berger AH, et al. : Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150** : 1107-1120, 2012. doi : 10.1016/j.cell.2012.08.029.
- 11) Pleasance ED, Stephens PJ, et al. : A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463** : 184-190, 2010. doi : 10.1038/nature08629.
- 12) Hodis E, Watson IR, et al. : A landscape of driver mutations in melanoma. *Cell* **150** : 251-263, 2012. doi : 10.1016/j.cell.2012.06.024.
- 13) Greenman C, Stephens P, et al. : Patterns of somatic mutation in human cancer genomes. *Nature* **446** : 153-158, 2007. doi : 10.1038/nature05610.
- 14) Nik-Zainal S, Alexandrov LB, et al. : Mutational processes molding the genomes of 21 breast cancers.

- Cell **149** : 979-993, 2012. doi : 10.1016/j.cell.2012.04.024.
- 15) Ellis MJ, Ding L, et al. : Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* **486** : 353-360, 2012. doi : 10.1038/nature11143.
- 16) Alexandrov LB, Nik-Zainal S, et al. : Signatures of mutational processes in human cancer. *Nature* **500** : 415-421, 2013. doi : 10.1038/nature12477.
- 17) The Catalogue Of Somatic Mutations In Cancer, <https://cancer.sanger.ac.uk/cosmic/signatures>
- 18) Pan-cancer analysis of whole genomes. *Nature* **578** : 82-93, 2020. doi : 10.1038/s41586-020-1969-6.
- 19) Alexandrov LB, Kim J, et al. : The repertoire of mutational signatures in human cancer. *Nature* **578** : 94-101, 2020. doi : 10.1038/s41586-020-1943-3.
- 20) Li Y, Roberts ND, et al. : Patterns of somatic structural variation in human cancer genomes. *Nature* **578** : 112-121, 2020. doi : 10.1038/s41586-019-1913-9.
- 21) Davies H, Morganella S, et al. : Whole-Genome Sequencing Reveals Breast Cancers with Mismatch Repair Deficiency. *Cancer Res* **77** : 4755-4762, 2017. doi : 10.1158/0008-5472.Can-17-1083.
- 22) Davies H, Glodzik D, et al. : HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nat Med* **23** : 517-525, 2017. doi : 10.1038/nm.4292.
- 23) Nik-Zainal S, Davies H, et al. : Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534** : 47-54, 2016. doi : 10.1038/nature17676.
- 24) Nik-Zainal S, Davies H, et al. : Author Correction : Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **566** : E1, 2019. doi : 10.1038/s41586-019-0883-2.