基礎老化研究 46（3）; 35－45, 2022

【総説】

診断バイオマーカーとしてのDNAメチル化

新飯田　俊平、下田　修義
国立研究開発法人国立長寿医療研究センター研究所
研究推進基盤センター

要約

医療現場において血中DNAメチル化が疾患の診断マーカーとして利用されるのは、メチル化変動の明瞭なインプリンティング異常による疾患にほぼ限られていた。しかしながら、DNAのCpGという並びにあるシトシン塩基のメチル化修飾は共有結合していることから非常に安定しており、疾患部位のメチル化変動が血液でも検出されれば、他の疾患でもサロゲートマーカーになりえる。実際、次世代シーケンサーの登場でメチル化DNA解析能力が飛躍的に向上したことにより、イギリスではがん患者の血液にわずかに含まれる遊離DNAのメチル化をもとに、多種類のがんを超早期に発見しようとする大規模臨床研究が始まっている。血液DNAメチル化は、認知症をはじめとするさまざまな多因子疾患のための新たな早期診断マーカーとして期待が持たれている。

キーワード：DNA methylation, cfDNA, ctDNA, complex disease, EWAS

1. はじめに

11年前、筆者の一人は本学会誌に、老化の理解にはモデル生物における実験的な介入による寿命の伸び縮みではなく、ヒトの高齢者特有の疾患の予測実現が有効ではないかと記した[1]。その後の十年で加齢に伴って変動するDNAメチル化は、脊椎動物亜門の様々な動物において年齢をかなり正確に推定できる標的になることがわかってきた[2]。この年齢予測は再現性の高さから、老化研究分野以外においても活用されはじめている[3, 4]。

加齢の終盤に避けがたく生じるのが老化である。これを病的なものととらえるかは意見の分かれるところであるが、老化そのものが少なからず高齢者疾患の下地を作っていることは確かである（predispose）。がんの発症率は年齢と共に高まる。認知症に至っては加齢が最大のリスクファクターである。仮に老化を前病的なものと見なすなら、DNAメチル化による年齢測定の成功は高齢者疾患の前兆を予見するツールになることを示唆する。

連絡先：新飯田　俊平
〒474-8522　愛知県大府市森岡町7丁目430番地
TEL：0562-44-5651
E-mail：sniida@ncgg.go.jp

一方でDNAメチル化は、その物質としての安定性から、明確なメチル化変化を伴う疾患ではすでにバイオマーカーとして利用されている。高齢者疾患の発症予測にDNAメチル化を利用しようと考えるのは必然的である。あるいは、疾患の発症予測とまでいかずとも、たとえばマウスを利用した疾患発症機序研究にDNAメチル化の視点を取り入れたいと考えても不思議はない。本稿ではそのような興味を持つ方に少しでも参考になればと、診断マーカーとしてのDNAメチル化について、1）DNAメチル化が診断マーカーとしてすでに利用されているインプリンティング疾患、2）バイオマーカーとして大きな進展の期待されるがん、3）利用が期待されている多因子疾患、という三つの疾患領域に即して紹介する。

2. インプリンティング異常疾患の診断

難病であるインプリンティング疾患（表1）は、現在、DNAメチル化が補助的あるいは確定的な診断マーカーとして臨床現場で利用されている、おそらく唯一の疾患領域である。

インプリンティングについて簡単に説明してみたい。父母由来の一対の染色体の相同な場所に位置し、機能的にも相同な遺伝子を対立遺伝子（アレル）という。通常一対のアレルは等しい転写活性を有するが、哺乳類（ただ

表１　DNA メチル化を診断マーカーとする主なインプリンティング疾患

疾患名（英語の略称）	OMIM	発症頻度*	検査方法	国内検査実施施設
アンジェルマン症候群 (AS)	#105830	1/12,000~1/20,000	MS-PCR パイロシーケンス	SRL、LSI 佐賀医、成育医療
プラダー・ウィリ症候群 (PWS)	#176270	1/10,000~1/30,000	MS-PCR パイロシーケンス	SRL、LSI 佐賀医、成育医療
ラッセル・シルバー症候群 (RSS)	#180860	1/100,000	パイロシーケンス	佐賀医、成育医療
ベックウィズ・ヴィーデマン症候群 (BWS)	#130650	国内推定 200 人以上	パイロシーケンス	佐賀医、成育医療

* 発症頻度は GENE Reviews Japan（GRJ）による（BWS を除く）。佐賀医：佐賀大学医学部、成育医療：国立成育医療センター、LSI: LSI メディエンス

し単孔類は除く）では父親由来遺伝子と母親由来遺伝子のうち、決まってどちらかの親に由来したアレルのみが発現する遺伝子があり、それをインプリンティング（刷り込み）遺伝子と呼んでいる。インプリンティング遺伝子は染色体上の複数箇所にまとまって存在し、それぞれドメイン（インプリンティングドメイン）を形成する。インプリンティングドメインと同じ遺伝子座には CpG アイランドか CpG に富んだ配列からなる「インプリンティングセンター」がある。ここでは父性染色体と母性染色体で異なる DNA メチル化を示すメチル化差異領域（Differentially Methylated Region; DMR）を形成する。インプリンティングセンターがメチル化されているか、逆にメチル化を免れているかで *cis* に存在するインプリンティングドメイン内の個々のインプリンティング遺伝子がオフになるかオンになるかが決まる[5]。なお、インプリンティング遺伝子は父性発現型と母性発現型がほぼ同数存在する。DNA メチル化がインプリンティングに重要な役割を果たしていることは、DNA メチル化を欠損したマウス胎児ではインプリンティング異常が見られることからも明らかである[6, 7]。

インプリンティング疾患はインプリント遺伝子の発現異常により引き起こされる。その原因は以下の四つに求められる。１）インプリンティングセンターのメチル化異常、２）染色体不分離などの原因により、一対の相同染色体が丸ごと、あるいは部分的（partial or segmental）に片親由来になってしまう片親性ダイソミー（uniparental disomy; UPD）、３）発現すべき父性あるいは母性アレルの欠損、４）インプリンティング遺伝子の変異。１）～３）は図２からイメージしていただきたい。代表的なインプリンティング疾患を表１に示す。それぞれの疾患の詳しい臨床像、発症機序等については OMIM を、またメチル化解析法の原理については下田の総説を参照いただければ幸いである[1]。

DNA メチル化が診断に使われるわかりやすい例として、プラダー・ウィリ症候群（Prader-Willi Syndrome; PWS）とアンジェルマン症候群（Angelman Syndrome; AS）について紹介する。プラダー・ウィリ症候群は胎児期の胎動減少、乳児早期の重度の筋緊張低下と摂食障害、乳児期後期から小児早期に始まる過食、徐々に進行する病的肥満症と中程度の発達障害を特徴とする。一方、アンジェルマン症候群は重度の運動・言語発達障害、精神遅滞、特異な行動などを特徴とする。実はこれら二つの異なる疾患は 15 番染色体上の 15q11.2-q13 に存在する同じインプリンティングドメインの異常に起因する[5]。そのインプリンティングドメインで、父由来の染色体で発現すべき遺伝子が発現できないとプラダー・ウィリ症候群となり、対照的に母親由来の染色体で発現すべき遺伝子が発現できないとアンジェルマン症候群となる（図２）。

検査に使われるのはメチル化を定性的もしくは定量的に解析する手法である（表１）[1]。たとえば検査会社の SRL では 15q11.2 領域に存在する *SNRPN* 遺伝子（small nuclear riboprotein-associated peptide N）の 5’ CpG アイランドを対象としたメチル化特異的（Methylation Specific; MS）PCR 法を採用している（図１）[8]。ちなみに *SNRPN* 遺伝子の CpG アイランドこそが上述のインプリンティングセンターに相当する。この CpG アイランドは父由来アレルではメチル化されていないのに対し、母由来アレルではメチル化されている。そのためメチル化状態の *SNRPN* にのみアニールできるプライマーと、それとは逆に非メチル化状態の *SNRPN* にのみアニールできるプライマーで PCR をかけると、前者は父由来のアレルのみを、後者は母由来のアレルのみを区別して増やすことができる（図２）。これを利用すれば 15q11.2-q13 の１）片側アレルの欠損、２）UPD、３）インプリンティング異常について、父アレルの異常であるのか、母アレルの異常であるのか見分けることができる。

実際のところプラダー・ウィリ症候群やアンジェルマン症候群の場合、１）、２）で発症原因の８割近くを占めるが、MS-PCR を使えば１）、２）の染色体構造異常だけでなく、３）のメチル化異常も確認できるので、DNA メチル化が有効なバイオマーカーになり得ている。

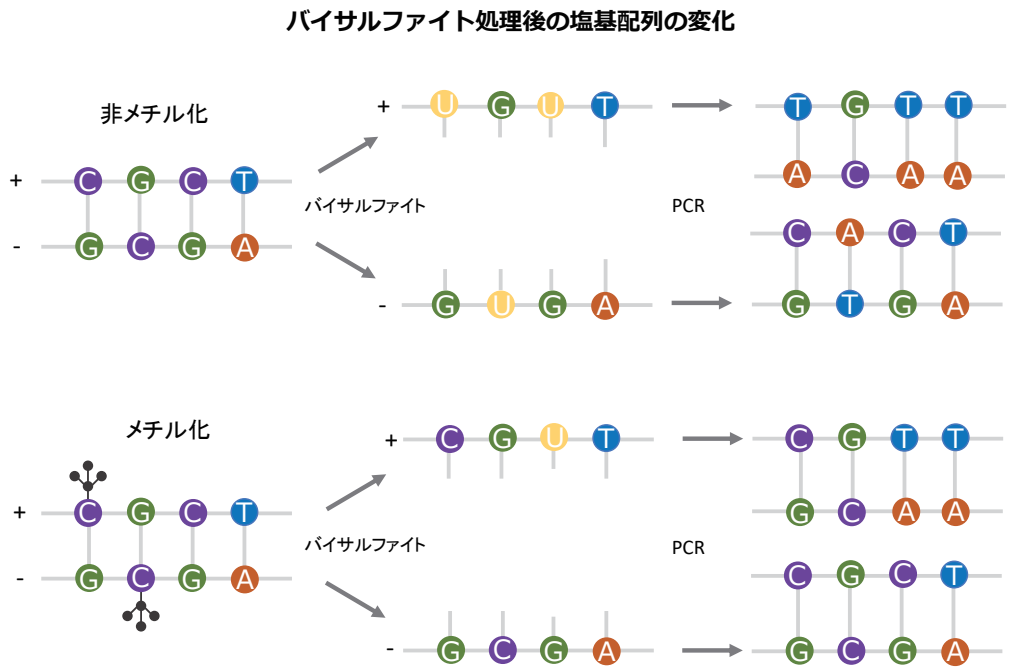

図１　メチル化 DNA 解析の基本

DNAを一本鎖に変換後、バイサルファイト処理するとメチル化されていないシトシンはウラシルに変わる。一方、メチル化されたシトシン(直後はグアニン）はシトシンのままである[46]。ここで DNAの +鎖と－鎖はもはや相補的ではなくなる。通常の DNAメチル化解析では、バイサルファイト処理後、CpG配列を含まない領域に対してデザインしたプライマーで PCRをかける。シーケンスして、その産物に“CpG” という配列が存在すれば、そのシトシンはメチル化されていたことを表す。メチル化特異的 PCRでは逆に、CpGが複数存在する領域に PCRプライマーを設定する。このときプライマーは非メチル化 DNAの配列(+鎖でも、－鎖でもよい）に合わせた配列と、メチル化 DNAの配列(+鎖でも、－鎖でもよい）に合わせた配列の２種類用意する。この両者のプライマーで個々に PCRを行えば、標的配列のメチル化･非メチル化を PCR産物の有無に変換し検出することができる。

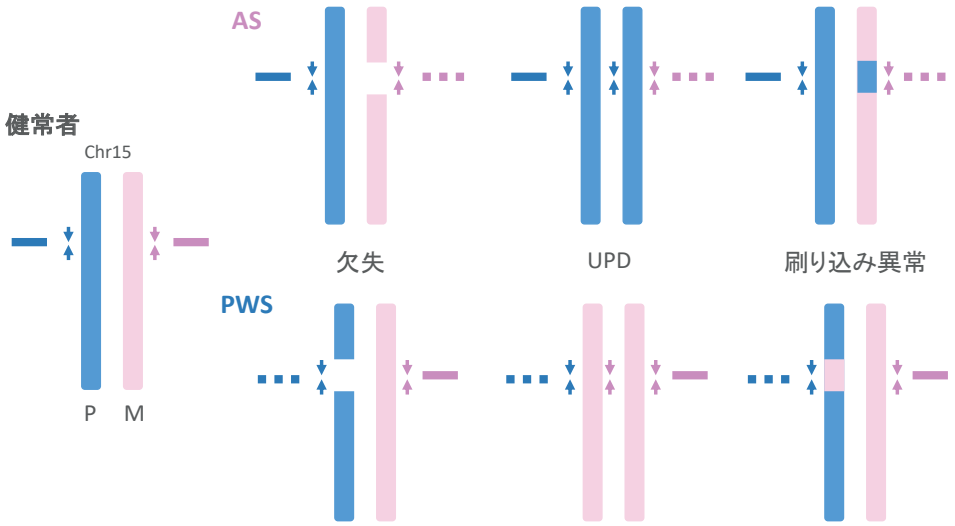

図２　メチル化特異的 PCR によるインプリンティング疾患の診断の例

向かい合う矢印がインプリンティングドメインに存在する *SNRPN* 遺伝子の CpG アイランドに対するプライマーセット。ここでは青色が非メチル化アレルに、ピンク色がメチル化アレルにハイブリできる。横向きの短い実線が PCR 産物のイメージで、点線は産物ができなかったことを表す。アンジェルマン症候群（AS）では、母親由来のメチル化した *SNRPN* 遺伝子に由来する PCR 産物が得られない。プラダー・ウィリ症候群（PWS）ではその逆となる。UPD: uniparental disomy

3. ctDNA によるがんの診断

これまで胎児の染色体異常(異数性)のスクリーニングでは母体に負担のかかる侵襲性の羊水穿刺による出生前検査が行われてきた。現在ではリキッドバイオプシー(血液)による低侵襲性の出生前検査が一般的になった。がんのサブタイピングと診断は、病理組織標本に基づく解析が昔も今もゴールドスタンダードである。そこから治療法の選択、予後予測、がんの原発巣やがんの分類などについて貴重な情報を得る。ただしこの伝統的な手法には、侵襲的な生検(バイオプシー)が必要となる。最近は血漿の中に存在する DNA でがんを早期発見しようとする取り組みが盛んである。この方法ならがんのモニターを *in situ* でできる。うまくいけば低侵撃性のリキッドバイオプシー検査で治療法の選択、がん細胞の残存・再発のモニターも可能になる[9]。

特に注目されているのはさまざまな組織の細胞から放出され、血液や尿の中を自由に漂う断片化した DNA である(図 3)。それらは無細胞核酸、cell-free DNA (cfDNA)と総称され、血液の場合、主に造血幹細胞に由来するが、そればかりではなく、一部は血管内皮細胞や肝細胞にも由来する[10]。がんが存在する場合、cfDNA にはがん細胞の細胞死に由来する cfDNA も含まれる。それを特に「血中循環腫瘍 DNA(circulating tumor DNA: ctDNA)」と呼ぶ[11]。たとえば健常者の血清には cfDNA はほとんど含まれていないが(1 ミリリットル当たり 0 ～ 50 ナノグラム、平均 13 ナノグラム)、がん患者ではその十数倍の平均 180 ナノグラムほどになる[12]。放射線治療でその量が減ることから、がん患者での cfDNA の増加分は ctDNA ではないかと推測されている。核に由来する健常者の cfDNA の最頻サイズはおよそ 166 bp で[13]、がんに由来する最頻 cfDNA (ctDNA)サイズはおよそ 143 bp とやや小さい。いずれもゲノムが断片化したものであることに変わりはない。DNA のヌクレオソーム一巻きがおよそ 147 bp で、これにリンカー断片がつくとおよそ 167bp になることから[14]、cfDNA は主にアポトーシスの結果から生じるのではないかと想像されているが[15]、それ以外にもネクローシス[16] やヌクレオプロテインとして細胞から積極的に排出される[17] という経路も示唆されている。

発がんにおいて、DNA メチル化の変化はもっとも初期かつもっとも広範に生じるゲノム変化と考えられている[18, 19]。正常な脊椎動物の細胞では CpG アイランドは非メチル化されており、繰り返し配列を含むそれ以外の領域は高度にメチル化されている。しかし、がん細胞のゲノム(以下、がんゲノム)のメチル化パターンはちょうどそれと反転するイメージである(図 4)[20]。がんゲノムでは、巨視的には繰り返し配列を主とするゲノムの全体的な低メチル化と、逆に、多くの CpG アイランドを含む局所的な高度メチル化が同時発生している[21]。がん化によりメチル化される CpG アイランドには主要ながん抑制遺伝子のものが含まれるため[22, 23]、がん抑制遺伝子のサイレンシングが起こる[24]。繰り返し配列の低メチル化はゲノムの不安定性をもたらすなど、それ自体に発がん性リスクがある。そのため、体細胞突然変異

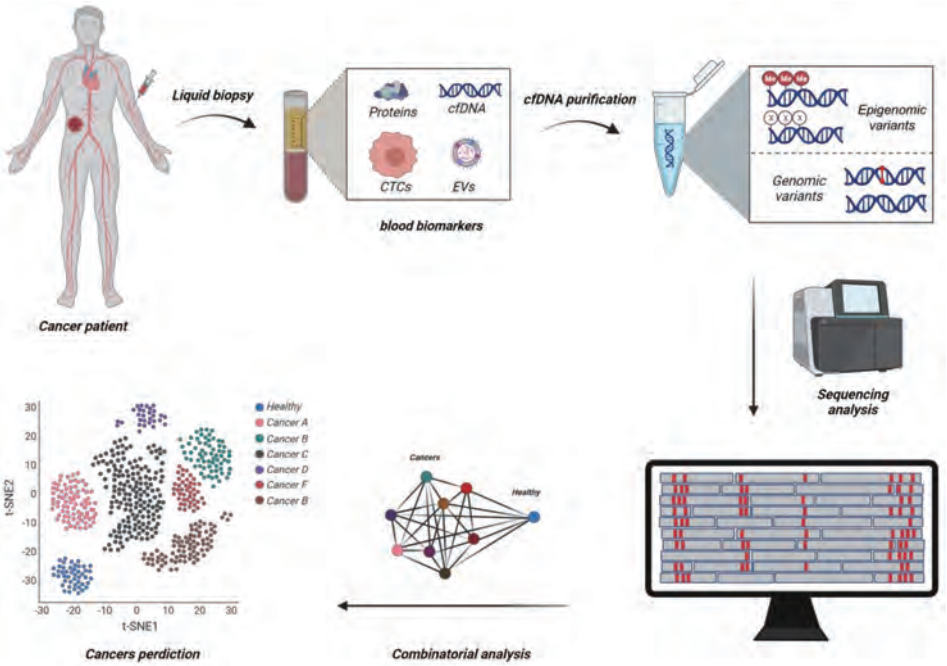

図3　リキッドバイオプシーによる多種のがんの同時検査

血液中の無細胞核酸(cfDNA) の変異や DNAメチル化から、さまざまながんを超早期に発見しようとする試みがなされている。DNAと異なりメチロームは組織ごとに異なる。そのため cfDNAのメチローム解析からは異常メチル化を受ける遺伝子の情報のみならず、がんの原発部位の情報も含まれる。あらかじめそれぞれのがん患者の cfDNA メチル化パターンをデータベース化しておけば、未症状のがん患者の血液からがんの種類を推定できる。図は[50] からの引用で BioRender.comが作製。CTCs: circulating tumor cells, EVs: extracellular vehicles

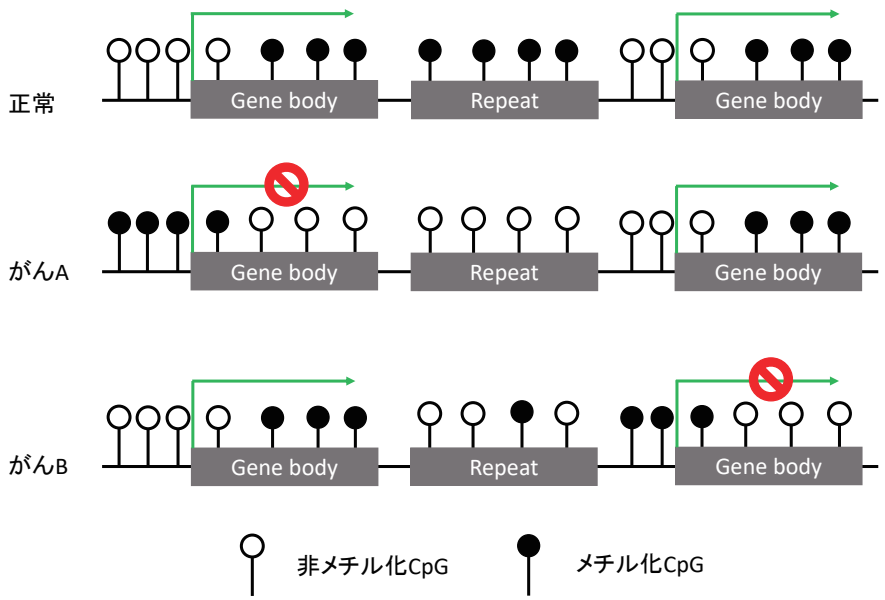

図４　がん細胞のゲノムで見られる DNA メチル化異常のイメージ
　がんの種類によってメチル化を受けやすい遺伝子のグループが存在する。その情報を利用した、特定のがんの有無を検査するキットが販売されている。

によるがんのドライバー変異にこれらのメチル化異常が組み合わさることで、がんの発生が容易になる[25]。

がんゲノムに生じる DNA メチル化の変化は微視的には多様かつ複雑であるが、その変化は多くのがんで組織を越えて、あるいは原発巣において高度に一致している。この特徴と DNA メチル化の物質としての安定性はがんの診断開発のための魅力的な特性であり、すでに血漿や便、FFPE サンプルからの DNA でがんを検出するための検査キットが上市されている[26]。上述のように、がん細胞ではがん抑制遺伝子やがん種に応じた固有遺伝子の体細胞突然変異、あるいは繰り返し配列の不安定性などが見られ、当然ながら ctDNA も同じ変異を有する。そのため cfDNA の網羅的な配列解析から ctDNA の含有、すなわちがんの有無を知ることができ、その変異部位(遺伝子)からがんのタイプがある程度推定できる。そこで複数の企業がリキッドバイオプシーで得られる cfDNA を用いたがん早期診断技術の開発を行っている。たとえば、イルミナ社からは 500 を超えるがん関連遺伝子の体細胞変異を一度に ctDNA から検出するキットが販売されている。

一方、前項に記したようにがん細胞ではそれぞれのがんに特徴的なメチル化の異常も生じ、それも ctDNA に保存される。このとき往々にしてメチル化の変化は単独の CpG 部位ではなくドメインとしてひとかたまりの CpG 部位に及ぶため、異常なメチル化のハプロタイプブロックを形成する[27]。これは点突然変異などの体細胞突然変異よりも正常型と区別しやすい。またゲノムには組織特有のメチル化パターンを持つ領域が存在する。そのため網羅的に cfDNA のメチル化解析をすると cfDNA の由来組織とその割合が計算できる(いわゆるデコンボリューション)。たとえば、ある組織のメチル化パターンをもつ cfDNA の割合が全 cfDNA 中で異常に増えたとすれば、それはその組織のがん化とそれに伴うアポトーシスにより組織からの cfDNA 流出が増加したことを示唆する。これは通常の DNA 配列からは得られない情報である[10]。そしてそれらの情報を組み合わせることで、がんの有無と原発巣が推定できることになる[28]。体内にがん細胞が見つかるものの、その細胞が増殖を開始した場所(原発部位)を特定できないがん、いわゆる原発不明がん(cancer/carcinoma of unknown primary; CUP)についても cfDNA により原発巣が明らかになる可能性がある。

cfDNA の変化を見る方法は、それが体細胞突然変異にしろ、メチル化の変化にしろ、血液検査で済むので繰り返し検査にそれほどの抵抗がないことから、がんの集団検診に向いている。多種のがんを早期に発見することは治療の機会を増やし、死亡リスクを下げるために極めて重要である。米国では、多種のがんに対する早期発見検査法の開発を目指した「循環無細胞ゲノム地図帳研究」The Circulating Cell-free Genome Atlas study(CCGA)という症例対照観察研究コンソーシアムが組織され、ctDNA の体細胞変異解析とメチル化解析が比較された。その結果、がん発見能力においては、メチル化が上述の ctDNA の体細胞変異解析を遙かにしのぐことが判明した。アメリカのスタートアップ企業で「聖杯」の意味をもつグレイル(GRAIL)社(イルミナのスピンオフ企業で再度イルミナが買収)は、このようながん発見における DNA メチル化の優位性を活かし、ctDNA のメチル化の特徴(シグニチャー)からがんの存在のみならず、その原発部位までを捕らえられることを可能にしたサービス、ガレリ(Galleri)の提供を始めた。インターネットの情報によると、すでに 2021 年には全米で医者の判断でガレリを利用できるようになっている。ただし FDA の承認はされておらず、約 1,000 ドル(13 万円)の費用がかかる。これについてイルミナは、メディケイドによる保

険適用がなされるよう望んでいる。ガレリはCCGAコンソーシアムのサブコホートでテストされ、一般的なスクリーニング検査の手法がない固形腫瘍(食道がん、肝臓がん、膵臓がんを含む50を超える種類)の検出を可能にするらしい[29, 30]。実はそれ以前に他のグループにより、ガレリとは別の免疫沈降を用いた手法で、ctDNAのメチル化シグニチャーにより複数のがんを見分けられることがNature誌に報告されていた[31]。こちらの方はその後の展開についての報告はない。一方、CCGA研究は、15,000人以上からなる前向きのコホートの構築と、そこから得られる血液およびがん組織の収集、機械学習を利用したがんの分類指標の設定など、領域横断的で圧倒的に豊富なデータ量とデータの質の高さがあり、それは結果としてガレリの有効性にお墨付きを与えることになった(ただしCCGA研究はガレリ販売元のグレイル社からサポートを受けており、論文著者の多くはグレイルとの関係がある)。そこでイギリスの国民保険サービス、National Health Service (NHS)は試験的に、ガレリを用いた165,000人を対象とする大規模な血液テストを2021年からスタートした(https://www.nhs-galleri.org/)。ここには14,000人からなる50歳から79歳まで未症状の人が含まれ、3年にわたり毎年ガレリ検査を実施する。この結果は2023年に報告される予定で、もし結果が良好であれば2年延長し、参加者をおよそ100万人まで拡大する計画である。そしてNHSは、治療できる可能性が高いステージ1～2でのがんの発見を、現在の50%から2028年までに75%へと高めることを目標にしている。この結果次第では、cfDNAのメチル化解析ががん検診におけるゲームチェンジャーになると予想されている。なおctDNAは単独で存在するわけではなく、他のcfDNAと混在するため早いステージのがんほどctDNAの量が少ない。したがってガレリであっても検出感度は落ちる。たとえば、死因となるがんのうち過半数を占める12のがん種においてステージ別にみると、ステージIVでは90.1%、ステージIIIで77.0%だが、ステージIIでは40.4%、ステージIで16.8%となる[29]。ただしこの時の特異度は99.5%と疑陽性の割合は極めて低く、また原発部位予測の正解率はがん全般で88.7%と高い。今後cfDNAのメチル化解析技術の向上により、より感度が高まることが期待される。

4. ctDNAによる複合疾患の診断

cfDNAが細胞死に伴い流出する核酸であるならば、神経変性疾患では神経細胞に由来するcfDNAが増加するのであろうか？もしそうであれば、がんのケースと同じように、中枢神経(CNS)由来のcfDNAはそのメチル化パターンから判断できるはずである。加えて、異常なメチル化パターンを持つ遺伝子はCNS由来である可能性が高く、そこから診断マーカーや発症機序のメカニズムの手がかりが得られるかもしれない。

それを最初に検討したのがLehmann-Wermanらの研究で、彼らはcfDNAメチロームから由来組織をデコンボリューションできることを利用して、1)(神経の)多発性硬化症患者でのオリゴデンドロサイト由来cfDNAの増加、2)心停止による脳障害での脳由来cfDNAの増加、3)外傷性能損傷での脳由来cfDNAの増加、の三つの知見を得た[32]。さらにそれぞれの疾患において、異常なDNAメチル化を示す遺伝子や領域を見出した。また、LIMホメオボックス2(*LHX2*)という遺伝子のCpGアイランドは脳で特異的にメチル化を受けることが知られている。実際、アルツハイマー病におけるcfDNAに神経細胞から浸出してきたと推定されるメチル化を受けた*LHX2*のCpGアイランドが含まれていた[33, 34]。比較的最近では、神経とグリア由来のcfDNAを血漿から同定するなど細胞種の分離も可能であることが示された[35]。これらの結果は、未だ病因の理解が進んでいない神経変性疾患、とりわけ認知症発症のメカニズム解明におけるcfDNAメチローム利用の有効性に期待を持たせ、異常なメチル化を示す遺伝子がすでにいくつか報告されている[33, 34]。

5. 血中DNAによる複合疾患の診断

血中のDNAメチル化を診断マーカーに利用しようとする研究はcfDNAではなく、主に白血球に由来する血球DNAを利用して始まった。対象は多因子疾患である[36, 37]。現在、様々な多因子疾患がエピゲノムワイド関連解析(epigenome-wide association study; EWAS,「イーバス」と発音)と称される研究手法の対象になっている。多因子疾患とは、その名の通り、疾患をもたらすそれぞれ重みの異なる複数の要因の積み重なりが、ある閾値を超えた時に発症すると考えられる疾患である。代表的なものとして認知症、糖尿病、自己免疫疾患、アテローム動脈硬化性疾患、神経障害や精神障害などがある。他にもさまざまな疾患がこのカテゴリーに分類される。

多因子疾患の診断発症の要因は、遺伝子と環境に大別できる。環境には、食事や運動、学習というある程度自分の意思でコントロールできる生活習慣一般から、大気汚染や心的障害、幼児期における母親からの愛情といった不可避的で数値化の困難な経験まで、様々な要因を含む。これらの環境によりDNAメチル化が影響を受けることが明らかにされつつある。そのため、それぞれの疾患に対し、発症に寄与するメチル化異常の検出を求め、これまで多数のEWASが行われてきた。しかし残念ながら、それらの研究からマーカーの上市に至った例を筆者らは知らない。そもそも筆者らの関心がある認知症に限っていえば、再現性の確認されたEWASの結果～具体的にはメチル化変動の見られた遺伝子～というものはほとんど報告されていない。せいぜいアンキリン1(*ANK1*)[38, 39]、ソルビンとSH3ドメイン含有3(*SORBS3*)[40, 41]、そしてヒストン脱アセチル化酵素2(*HDAC2*)[42, 43]くらいではなかろうか。一方、候補遺伝子アプローチにより、アルツハイマー病関連遺伝子*TREM2*がアルツハイマー病患者の血中で高メチル化傾向にあると報告されたが[44]、筆者らはその再現性を確認できた[45]。したがって認知症でのメチル化異常は確かに存在する。しかし現状のEWASではその検出に不確実性があるように見受けられる。このままでは、DNAメチル化異常をバイオマーカーに、あ

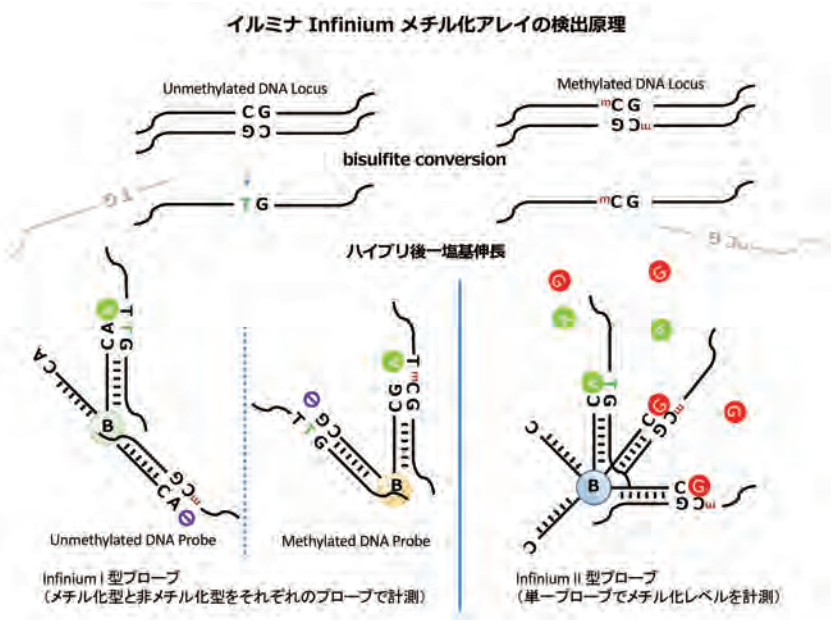

図５　イルミナ社のアレイによるメチル化測定原理と問題点
　イルミナのメチル化アレイには２種類のプローブが搭載されており、どちらのタイプもビーズ（図中 Ⓑ）に個々のプローブが多数結合している。バイサルファイト処理後の一本鎖 DNAをプローブにハイブリダイズさせた後、プローブから蛍光物質の付加された塩基を一つ取り込ませる。注意すべきはターゲットのシトシンが多型によりチミンに置き換わっていると（図６）、それは非メチル化シトシンとして測定されてしまうことである。またターゲット CpG周辺の、つまりプローブの 3' 末端側に存在する多型もハイブリに影響を与え、メチル化レベルの測定を誤らせる要因となる。

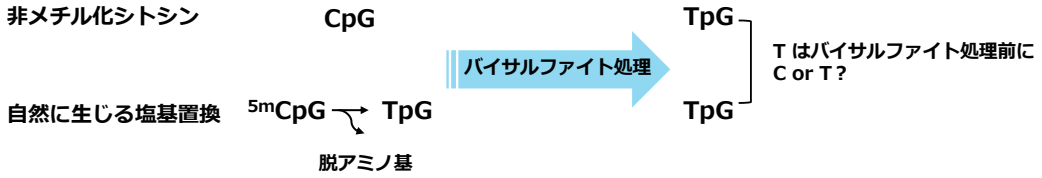

図６　多型がもたらすメチル化レベルの測定誤差
　バイサルファイト処理後の TpG配列のチミンは、元は非メチル化 CpGだったのか、多型によりはじめから TpG だったのか判別できない。イルミナ社のアレイプローブは CpG/TpG多型部位には設計されていないという前提だが、そうでない場合、TpG多型部位は非メチル型 CpG部位と誤認されることになる。

るいは発症機序につなげるという研究は、EWAS の再現性の低さから疑問視されてしまうことも危惧される。

　なぜこのような状況になったのか？筆者らの見立てでは EWAS でほぼ独占的に用いられているプラットフォームであるイルミナ社のメチル化解析用ビーズアレイそのものの問題（図５）と、研究者によるその利用の仕方に問題があるのではないかと考えている。そこで最後にこのアレイの問題点を提起したいくつかの報告をまとめてみたい。メチル化アレイの使用を検討している方々の参考になれば幸いである。

　イルミナ社はヒトゲノムのメチル化解析用のアレイを 2009 年から販売している。2016 年に販売された "Infinium MethylationEPIC BeadChip"（以下、メチル化アレイ）は三代目となる。検出原理については図５および参考文献１を参照していただきたい。このメチル化アレイでは、ヒトゲノムに存在する 85 万箇所の CpG 部位のメチル化レベルを一度に測定できるという説明になっている。それらの CpG 部位はほぼすべての遺伝子をカバーしていて、それらの遺伝子の内部にはサーベイされる CpG 部位が少なくともいくつかは存在する。もっともヒトゲノムには CpG 配列が 2,800 万箇所存在するので、測定される CpG 部位の割合としては全体の 3% にすぎない。また、それらの大半は遺伝子のプロモーター領域に偏重する傾向にあり、ゲノムワイドに敷設されたプローブというイメージからは実はほど遠い。最善の手法は全ゲノムのバイサルファイトシーケンスであるが、多サンプルに対してはコストの面から見て現実的ではない。そのため現在 EWAS といえば、比較的廉価に行え、受託業者も複数登場したメチル化アレイ解析のことを指すといってよい。

　そのメチル化アレイの深刻な問題は、アレイに搭載されたプローブが不可避的に DNA 多型をメチル化の差と誤認してしまうことである（図６）。原因は以下になる。メチル化シトシン（通常 CpG サイトのシトシン）は自然

の脱アミノ化によりチミンへと変化する。このC(pG)からT(pG)へのトランジションは最も頻度の高い多型生成で、当然、同一の民族でも、ましてや異なる民族間ではより頻繁に見られることになる。一方、日本人研究者、早津が開発し、現代のメチル化解析において欠かせない技術となったバイサルファイト法はバイサルファイト処理により非メチル化シトシンをチミン(正確にはウラシル)に変化させることで、同処理により変化しないメチル化シトシンと区別する(図1)[46]。さて、もしここであるCpGサイトが、メチル化の測定対象となっているときに、メチル化アレイはバイサルファイト処理後のサンプルを利用するので、そこがTpGであると、そのCpGサイトのシトシンは非メチル化シトシンとして認識する。しかしそのTpGはそもそもバイサルファイト処理前から多型によってTpGだったかもしれないのである(図6)。このときサンプル集団内でのTのアレル頻度が高いほど、見かけ上、低メチル化に振れてしまうことになる。ここで二つのグループ間でのメチル化差を比較すると、その二グループ間でのアレル頻度の差に応じて、偽のメチル化差となって提示される。もっとも、多型が原因のメチル化異常はサンプルのメチル化レベルをプロットするとガウス分布ではない、特徴的な分布を示すので[47]、メチル化マーカーの候補が得られたらそのメチル化レベルのサンプル集団における分布を視覚化することが大切である。実際、明らかに多型によると見られるメチル化差のプロットが論文で提示されているケースがある[48]。ある民族間のメチル化差を調べた研究においては、異なる民族の間ということもあるが、メチル化アレイのプローブの実に7割がそのような多型部位に設置されていたという[49]。ちなみに、最近マウスの280Kメチル化チップが販売されたが、純系の利用できるマウスではもちろんこのような問題とは無縁のはずである。

イルミナはプローブ設計において多型部位は避けたとしているが、すべての多型部位を避けられるわけでもなく、またそうしようとすることも現実的ではない。要はユーザーが多型の問題の存在を認識して、提示されたメチル化異常のデータを視覚化するなど精査すると共に、メチル化異常部位候補については必ず他の手法を利用してバリデーションをすればよいのである。

イルミナ社は、3代目アレイに続いて、バイサルファイト処理後の標的領域を濃縮する技術(ツイストバイオサイエンス)を次世代シーケンサーと組み合わせたTruSeq Methyl Capture EPIC(メチルキャプチャー)を販売した(図7)。メチル化測定できるCpGサイトが330万箇所と現行アレイの4倍近くになった。当然、サンプル当たりのコストはそれ以上になった。しかし、シーケンスリードならばCpG/TpG多型部位があったとしても、その近傍のSNPからハプロタイプブロックとして認識できるのではないか。これはメチル化アレイに対するメチルキャプチャーの大きなアドバンテージとなる。次世代シーケンサーの性能向上に伴いデータ産出量あたりの単価は下落しており、もし多因子疾患に対するメチル化アレイの限界が見えてくるようになれば、EWASの主要なプラットフォームは今後このメチルキャプチャーが取って代わるかもしれない。

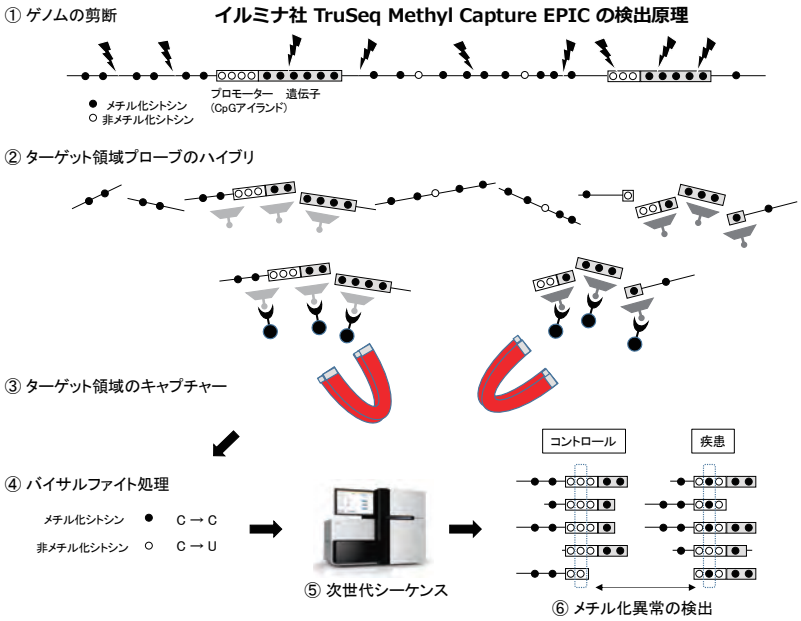

図7　ターゲット濃縮型メチル化解析法

ゲノムをランダムに剪断後、シーケンス用のアダプターを付加する。ついで、主に遺伝子に対して設計されたビオチン付のプローブとハイブリさせ、ストレプトアビジンとのアフィニティーを利用して標的配列の濃縮をかける。濃縮されたDNAに対しバイサルファイト処理する。ここで一本鎖となったDNAを鋳型としてPCR反応を行い、シーケンスの鋳型となるDNAライブラリーを作製する。そのライブラリーを両側から100 bpずつ読み、リードはイルミナ社のクラウド上に置く。同クラウド上で利用できるソフトウェアを用いてメチル化レベルの比較を行う。

6. まとめ

本稿ではDNAメチル化が診断マーカーとして、1）すでに利用されている、2）利用されつつある、3）利用が期待されている、三つの疾患領域について概説した。このうち、2）の体液（リキッドバイオプシー）に含まれる無細胞核酸（セルフリーDNA; cfDNA）のメチル化情報に基づくがん検診は近年特に注目を集めており、詳しく解説した。その理由は、cfDNAメチロームががんにとどまらず、将来的に認知症やフレイルを初めとする高齢者疾患の発症予測にも利用できるかもしれないという期待があるからである。前者は神経細胞死を後者は筋繊維の減少を伴うため、それぞれの細胞から特徴的なメチロームを持つcfDNAが流出している可能性がある。年齢のみならず、将来的にこれらの疾患もDNAメチル化から予測できるようになり、加えてDNAメチル化の加齢変化がもたらす遺伝子発現への影響が理解できるようになれば、老化という現象がより理解されるのではないか、と筆者らは期待する。したがって、高齢者疾患のマーカーをDNAメチル化に求め探索するということは、予防医学や先制医学への貢献はもちろんのことではあるが、筆者らからすれば、老化の根本原理を解明するという老化研究そのものなのである。

引用文献

1. 下田 修義. 哺乳動物の加齢に伴うDNAメチル化の変動. 基礎老化研究 35: 27-37, 2011.
2. Seale K, Horvath S, Teschendorff A, et al. Making sense of the ageing methylome. Nat Rev Genet, 2022.
3. Mertens J, Herdy JR, Traxler L, et al. Age-dependent instability of mature neuronal fate in induced neurons from Alzheimer's patients. Cell Stem Cell 28: 1533-1548.e1536, 2021.
4. Lu Y, Brommer B, Tian X, et al. Reprogramming to recover youthful epigenetic information and restore vision. Nature 588: 124-129, 2020.
5. Robertson KD. DNA methylation and human disease. Nat Rev Genet 6: 597-610, 2005.
6. Li E, Beard C, and Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 366: 362-365, 1993.
7. Howell CY, Bestor TH, Ding F, et al. Genomic imprinting disrupted by a maternal effect mutation in the *Dnmt1* gene. Cell 104: 829-838, 2001.
8. Kubota T, Das S, Christian SL, et al. Methylation-specific PCR simplifies imprinting analysis. Nat Genet 16: 16-17, 1997.
9. Locke WJ, Guanzon D, Ma C, et al. DNA methylation cancer biomarkers: translation to the clinic. Front Genet 10: 1150, 2019.
10. Moss J, Magenheim J, Neiman D, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. Nat Commun 9: 5068, 2018.
11. Lo YMD, Han DSC, Jiang P, et al. Epigenetics, fragmentomics, and topology of cell-free DNA in liquid biopsies. Science 372, 2021.
12. Leon SA, Shapiro B, Sklaroff DM, et al. Free DNA in the serum of cancer patients and the effect of therapy. Cancer Res 37: 646-650, 1977.
13. Snyder MW, Kircher M, Hill AJ, et al. Cell-free DNA comprises an in vivo nucleosome footprint that informs its tissues-of-origin. Cell 164: 57-68, 2016.
14. Heitzer E, Auinger L, and Speicher MR. Cell-Free DNA and apoptosis: how dead cells inform about the living. Trends Mol Med 26: 519-528, 2020.
15. Martins-Ferreira R, Leal BG, and Costa PP. The potential of circulating cell-free DNA methylation as an epilepsy biomarker. Front Cell Neurosci 16: 852151, 2022.
16. Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 61: 1659-1665, 2001.
17. Stroun M, Lyautey J, Lederrey C, et al. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. Clin Chim Acta 313: 139-142, 2001.
18. Alvarez H, Opalinska J, Zhou L, et al. Widespread hypomethylation occurs early and synergizes with gene amplification during esophageal carcinogenesis. PLoS Genet 7: e1001356, 2011.
19. Feinberg AP, Ohlsson R, and Henikoff S. The epigenetic progenitor origin of human cancer. Nat Rev Genet 7: 21-33, 2006.
20. Bergman Y and Cedar H. DNA methylation dynamics in health and disease. Nat Struct Mol Biol 20: 274-281, 2013.
21. Ross JP, Rand KN, and Molloy PL. Hypomethylation of repeated DNA sequences in cancer. Epigenomics 2: 245-269, 2010.
22. Wong IH, Lo YM, Zhang J, et al. Detection of aberrant *p16* methylation in the plasma and serum of liver cancer patients. Cancer Res 59: 71-73, 1999.
23. Esteller M, Sanchez-Cespedes M, Rosell R, et al. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res 59: 67-70, 1999.
24. Belinsky SA, Nikula KJ, Palmisano WA, et al. Aberrant methylation of $p16^{INK4a}$ is an early event in lung cancer and a potential biomarker for early diagnosis. Proc Natl Acad Sci U S A 95: 11891-

11896, 1998.
25. Tao Y, Kang B, Petkovich DA, et al. Aging-like spontaneous epigenetic silencing facilitates Wnt activation, stemness, and *Braf*^{V600E}-induced tumorigenesis. Cancer Cell 35: 315-328.e316, 2019.
26. Taryma-Leśniak O, Sokolowska KE, and Wojdacz TK. Correction to: Current status of development of methylation biomarkers for in vitro diagnostic IVD applications. Clin Epigenetics 12: 107, 2020.
27. Guo S, Diep D, Plongthongkum N, et al. Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA. Nat Genet 49: 635-642, 2017.
28. Feng H, Jin P, and Wu H. Disease prediction by cell-free DNA methylation. Brief Bioinform 20: 585-597, 2019.
29. Klein EA, Richards D, Cohn A, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. Ann Oncol 32: 1167-1177, 2021.
30. Liu MC, Oxnard GR, Klein EA, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol 31: 745-759, 2020.
31. Shen SY, Singhania R, Fehringer G, et al. Sensitive tumour detection and classification using plasma cell-free DNA methylomes. Nature 563: 579-583, 2018.
32. Lehmann-Werman R, Neiman D, Zemmour H, et al. Identification of tissue-specific cell death using methylation patterns of circulating DNA. Proc Natl Acad Sci U S A 113: E1826-1834, 2016.
33. Pai MC, Kuo YM, Wang IF, et al. The role of methylated circulating nucleic acids as a potential biomarker in Alzheimer's disease. Mol Neurobiol 56: 2440-2449, 2019.
34. Guemri J, Pierre-Jean M, Brohard S, et al. Methylated ccfDNA from plasma biomarkers of Alzheimer's disease using targeted bisulfite sequencing. Epigenomics 14: 451-468, 2022.
35. Chatterton Z, Mendelev N, Chen S, et al. Bisulfite amplicon sequencing can detect glia and neuron cell-free DNA in blood plasma. Front Mol Neurosci 14: 672614, 2021.
36. Rakyan VK, Down TA, Balding DJ, et al. Epigenome-wide association studies for common human diseases. Nat Rev Genet 12: 529-541, 2011.
37. Heyn H and Esteller M. DNA methylation profiling in the clinic: applications and challenges. Nat Rev Genet 13: 679-692, 2012.
38. Lunnon K, Smith R, Hannon E, et al. Methylomic profiling implicates cortical deregulation of *ANK*1 in Alzheimer's disease. Nat Neurosci 17: 1164-1170, 2014.
39. De Jager PL, Srivastava G, Lunnon K, et al. Alzheimer's disease: early alterations in brain DNA methylation at *ANK1*, *BIN1*, *RHBDF2* and other loci. Nat Neurosci 17: 1156-1163, 2014.
40. Siegmund KD, Connor CM, Campan M, et al. DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. PLoS One 2: e895, 2007.
41. Sanchez-Mut JV, Aso E, Panayotis N, et al. DNA methylation map of mouse and human brain identifies target genes in Alzheimer's disease. Brain 136: 3018-3027, 2013.
42. Gräff J, Rei D, Guan JS, et al. An epigenetic blockade of cognitive functions in the neurodegenerating brain. Nature 483: 222-226, 2012.
43. D'Addario C, Candia SB, Arosio B, et al. Transcriptional and epigenetic phenomena in peripheral blood cells of monozygotic twins discordant for Alzheimer's disease, a case report. J Neurol Sci 372: 211-216, 2017.
44. Ozaki Y, Yoshino Y, Yamazaki K, et al. DNA methylation changes at *TREM2* intron 1 and *TREM2* mRNA expression in patients with Alzheimer's disease. J Psychiatr Res 92: 74-80, 2017.
45. Mitsumori R, Sakaguchi K, Shigemizu D, et al. Lower DNA methylation levels in CpG island shores of *CR1*, *CLU*, and *PICALM* in the blood of Japanese Alzheimer's disease patients. PLoS One 15: e0239196, 2020.
46. Hayatsu H, Wataya Y, Kai K, et al. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. Biochemistry 9: 2858-2865, 1970.
47. LaBarre BA, Goncearenco A, Petrykowska HM, et al. MethylToSNP: identifying SNPs in Illumina DNA methylation array data. Epigenetics Chromatin 12: 79, 2019.
48. Li QS, Vasanthakumar A, Davis JW, et al. Association of peripheral blood DNA methylation level with Alzheimer's disease progression. Clin Epigenetics 13: 191, 2021.
49. Daca-Roszak P, Pfeifer A, Żebracka-Gala J, et al. Impact of SNPs on methylation readouts by Illumina Infinium HumanMethylation450 BeadChip Array: implications for comparative population studies. BMC Genomics 16: 1003, 2015.
50. Constantin N, Sina AA, Korbie D, et al. Opportunities for early cancer detection: the rise of ctDNA methylation-based pan-cancer screening technologies. Epigenomes 6, 2022.

DNA methylation as diagnostic biomarkers

Shumpei Niida, Ph.D., Nobuyoshi Shimoda, Ph.D.
Research Core Administration,
Research Institute, National Center for Geriatrics and Gerontology

Abstract

Methylation modifications of cytosine bases in the CpG sequence of DNA are covalently bound and very stable, and can be a good surrogate marker if disease-associated methylation variations are found not only at the site of disease but also in the blood. Until now, the use of DNA methylation as a diagnostic marker in blood has been almost exclusively limited to imprinting diseases with clear methylation changes. However, with the advent of next-generation sequencers, the ability to analyze methylated DNA has dramatically improved, and a large-scale clinical study has begun in the UK to detect many types of cancer at very early stages based on the methylation of a small amount of free DNA in the blood of cancer patients. Blood DNA methylation has also shown promise as an early diagnostic marker for a variety of complex diseases, including dementia.

Keywords：DNA methylation, cfDNA, ctDNA, complex disease, EWAS