Adjustment of cell-type composition minimizes systematic bias in blood DNA methylation profiles derived by DNA collection protocols

Yuh Shiwa1,2, Tsuyoshi Hachiya1,2, Ryoei Furukawa2, Hideki Ohmomo2, Kanako Ono2, Hisaaki Kudo3, Jun Hata4,5, Atsushi Hozawa6, Motoki Iwasaki7, Koichi Matsuda8, Naoko Minegishi3, Mamoru Satoh1,2,9,10, Kozo Tanno11, Taiki Yamaji7, Kenji Wakai12, Jiro Hitomi13,14, Yutaka Kiyohara15, Michiaki Kubo16, Hideo Tanaka17, Shoichiro Tsugane7, Masayuki Yamamoto18,19, Kenji Sobue20,21, Atsushi Shimizu2#

1 岩手医科大学 いわて東北メディカル・メガバンク機構 メガバンク・データ管理部門
2 岩手医科大学 いわて東北メディカル・メガバンク機構 生体情報解析部門
3 東北大学 東北メディカル・メガバンク機構 生体情報解析部門
4 九州大学 大学院医学研究院 病態機能内科学
5 九州大学 大学院医学研究院 附属総合コホートセンター
6 東北大学 東北メディカル・メガバンク機構 予防医学・疫学部門
7 国立がんセンター がん予防・検診研究センター 予防研究部
8 東京大学 医科学研究所ヒトゲノム解析センター シークエンス技術開発分野
9 岩手医科大学 いわて東北メディカル・メガバンク機構 地域連携・医療情報 ICT 部門
10 岩手医科大学 医歯薬総合研究所 生体情報解析部門
11 岩手医科大学 医学部 衛生学公衆衛生学講座
12 名古屋大学 大学院医学系研究科 予防医学
13 岩手医科大学 いわて東北メディカル・メガバンク機構 副機構長
14 岩手医科大学 医学部 解剖学講座
15 九州大学 大学院医学研究院 環境医学
16 理化学研究所 ゲノム医科学研究所 ゲノム解析技術開発チーム
17 愛知県がんセンター研究所 疫学・予防部
18 東北大学 東北メディカル・メガバンク機構 ゲノム解析部門
＜概要＞
コホート研究は医学研究のなかでもデータの精度・信頼度が高いため、近年の個別化予防・個別化医療の実現に必須の研究分野です。コホート研究では生体試料の数が精度に影響するため、できるだけ多くの参加者の方が協力いただく必要があります。日本では我々東北メディカル・メガバンク機構（東北 MM；東北大学東北 MM、岩手医科大学いわて東北 MM）の他、久山町研究（久山）、多目的コホート研究（JPHC）、日本多施設共同コホート研究（J-MICC）、バイオバンク・ジャパン（BBJ）をはじめとする、多くのコホート研究が進められています。

ゲノムコホート研究では収集した血液あるいは血液から抽出した DNA を保存していますが、採血から DNA 抽出までの手順（DNA 収集プロトコール）は各コホートで異なります。ゲノム多型解析では、DNA 収集プロトコールの違いが研究結果に与える影響は非常に小さいと考えられますが、いわて東北メディカル・メガバンク機構が研究対象としている DNA のメチル化は環境により変化することが知られているため、DNA 収集プロトコールの違いが研究結果に影響を与えることが考えられます。しかし、これまで DNA 収集プロトコールの違いが DNA メチル化解析の結果に与える影響について、十分な知見はありませんでした。そこで、我々の研究グループでは、今後のゲノムコホート研究において国内の各コホートが連携して DNA メチル化解析を実施するために、DNA 収集プロトコールの相違が DNA メチル化解析に与える影響について詳細な解析を行いました。

まず、先に挙げた久山、JPHC、J-MICC、BBJ の各コホートから DNA 収集プロトコールの情報を提供いただきました。東北 MM の手法を加え、類似した手法をまとめると 21 種類もの異なる DNA 収集プロトコールがあることがわかりました。そこで、採血管の種類、輸送温度や時間を比較検討し、4 つの代表的な DNA 収集プロトコールを選択しました。続いて、岩手医科大学内でボランティアの方に協力して頂き、採血後すぐに DNA を抽出したコントロール試料を加えた 5 種類の DNA 収集プロトコールを用いて血液から DNA を抽出しました。
メチル化マイクロアレイを用いたDNAメチル化解析の結果、同一個人・同一日に採取した血液でもDNA収集プロトコールが異なるとDNAメチル化に差が生じていることが示されました。この原因をさらに調べたところ、DNA収集プロトコールの相違により、血液に含まれる細胞の種類の分布（細胞組成）に変化が生じていることがわかりました。つまり、異なるコホートで収集した生体試料を用いてDNAメチル化解析を行うことは難しいということです。そこで、我々は測定したDNAメチル化情報から元の血液に含まれていた細胞組成を推定し、その情報を用いてDNA収集プロトコールによるDNAメチル化の差を補正する手法を開発しました。これにより、コホートによって異なるDNA収集プロトコールで集められた血液由来DNAを相互利用できるようになり、コホート同士が連携して統合解析が行えることが示されました。

＜まとめ＞
本研究において、DNA収集プロトコールの相違がDNAメチル化に影響を与えることを世界で初めて明らかにし、その影響を補正して解析を行う手法も同時に確立しました。今後、各コホートで集めた生体試料を用いたDNAメチル化解析に本研究の成果を応用し、病気と関連のあるDNAメチル化の変化ならびに病気の発症前に変化するDNAメチル化を見つける方法について研究を進め、個別化予防・個別化医療の実現を目指します。

*1コホート研究
コホート研究は、ある集団内における疾病の発生確率を遺伝的素因や生活習慣、環境の違い等で識別する医学研究の1つです。さらに、遺伝子（ゲノム）の配列の違いやオミックス情報を要因の一つとして加えて行うのがゲノムコホート研究です。

*2ゲノム多型解析
ゲノム配列をDNAマイクロアレイや次世代型シークエンサーを用いて決定し、個人毎に異なる配列を調べる手法。日本では東北MMが昨年1070人のゲノム多型を公開している。

*3DNAメチル化
DNAのメチル化とは、DNAの遺伝暗号であるA・T・G・Cの4文字（塩基）の中の主にC（シトシン）塩基にメチル基（-CH₃）が結合し、遺伝子の働きを調節する仕組みの一つです。DNAのメチル化の異常は、がんや生活習慣病など様々な疾患に関わっており、現在注目を集めている研究分野です。
異なる機関で集められた血液（生体試料）は、異なる方法で収集されていたため状態が異なり、DNAメチル化解析のために相互に利用することが困難でした。

IMMではこのDNAメチル化の差は、生体試料の中の細胞の組成が変化していることが原因であることを突き止め、バイオインフォマティクスの手法で補正する方法を確立しました。

この補正方法を使うことで日本のコホート研究で保存している検体を相互に利用できるようになり、今後のコホート連携を加速することができるようになりました。