

博士論文

早産児・低出生体重児における血液検体での
DNA メチル化解析・遺伝子発現解析に基づいた
エピジェネティック変化の探索

鹿嶋 晃平

目次

略語一覧	10
要旨	12
I. 序文	13
1. 研究の背景	13
1-1. 低出生体重児の2大分類—早産児と Small for Gestational Age (SGA) 児	13
1-2. 低出生体重児と Developmental Origin of Health and Diseases (DOHaD)、そしてエピジェネティクスとの関連	14
1-3. これまでの早産児・低出生体重児領域のエピゲノムワイド関連研究	16
2. 研究の目的	18
II. 方法	19
1. 研究対象の集団	19
2. 研究対象集団の臨床データの収集および算出	23
3. 単核球分離と DNA および RNA の抽出	23
4. DNA メチル化アレイにおける Quality Control とデータの前処理	25

5. アレイデータによる単核球分離後 buffy coat 内の細胞分画の推定	28
6. 遺伝子発現アレイとデータ前処理	28
7. 統計解析	30
7-1. 在胎週数ないし出生体重 SD スコアと関連のある臨床的特徴の検 討	31
7-2. 在胎週数ないし出生体重 SD スコアについてのエピゲノムワイド関 連解析(EWAS) およびパスウェイ解析	32
7-3. 在胎週数ないし出生体重 SD スコアについての遺伝子発現関連解析 およびパスウェイ解析	36
7-4. DNA メチル化と遺伝子発現の関連の確認	39
7-5. エピジェネティックメモリーの候補絞り込み	40
III. 結果	42
1. 在胎週数ないし出生体重 SD スコアと関連する臨床的特徴	45
2. 在胎週数ないし出生体重 SD スコアについてのエピゲノムワイド関連解析 (EWAS)およびパスウェイ解析	49
3. 在胎週数ないし出生体重 SD スコアについての遺伝子発現関連解析および パスウェイ解析	58

4. DNA メチル化と遺伝子発現の関連の確認	66
5. エピジェネティックメモリーの候補	69
IV. 考察	86
V. 結論	101
倫理面への配慮	102
謝辞	102
参考文献	104

図一覧

図 1. 研究参加者のリクルート、および DNA メチル化解析用サンプルのフィルタリングまでの流れ	20
図 2. 遺伝子発現アレイ解析用の RNA サンプルのフィルタリングの流れ	22
図 3. 単核球分離および DNA/RNA 抽出までの工程	24
図 4. DNA メチル化アレイにおける品質チェックとデータの前処理	26
図 5. DNA メチル化解析における Probe Filtering	26
図 6. 遺伝子発現アレイ解析とデータ前処理の流れ	28
図 7. 全体の解析の枠組み	31
図 8. 転写開始点(TSS)を軸とした CpG サイトの所在に関するゲノム領域分類	35
図 9. CpG アイランドを軸とした CpG サイトの所在に関する分類	36
図 10. EWAS で得られた在胎週数関連 CpG と出生体重 SD スコア関連 CpG を対応する遺伝子発現プローブにマッチング	37
図 11. DNA メチル化-遺伝子発現の直接比較のために、在胎週数関連 CpG サイトと対応する在胎週数関連発現プローブの組み合わせを作成	40
図 12. 在胎週数ないし出生体重 SD スコアと出生前変数の関係 (n=110, 臍帯	

血)	45
図 13. 在胎週数ないし出生体重 SD スコアについてのエピゲノムワイド関連解析(EWAS)(n = 110, 臍帯血)	50
図 14. DNA メチル化と在胎週数の関連度(p-value) についての Manhattan Plot (n = 110, 臍帯血)	52
図 15. DNA メチル化と出生体重 SD スコアの関連度(p-value) についての Manhattan Plot (n = 110, 臍帯血)	53
図 16. 在胎週数関連 CpG ないし出生体重 SD スコア関連 CpG の分布	57
図 17. 在胎週数ないし出生体重 SD スコアについての遺伝子発現関連解析 (n = 55, 臍帯血)	59
図 18. 在胎週数関連遺伝子発現プローブと対応する在胎週数関連 CpG の DNA メチル化－遺伝子発現間の関連についての検討 (n = 55, 臍帯血)	67
図 19. 在胎週数関連のエピジェネティックメモリーの候補 CpG (n = 47, 臍帯血および生後末梢血)	70
図 20. 出生体重 SD スコア関連のエピジェネティックメモリーの候補 CpG(n = 47, 臍帯血および生後末梢血)	72
図 21. 在胎週数関連のエピジェネティックメモリーの候補 CpG についての補	

足(n = 47, 臍帯血および生後末梢血) 73

図 22. 在胎週数関連エピジェネティックメモリーの候補 CpG サイトにおける
遺伝子発現 78

表一覧

表 1. DOHaD 仮説と Barker 仮説の関係	15
表 2. 本研究における臍帯血 110 サンプルの臨床データの分布	42
表 3. 出生前変数と在胎週数の関連－単変量回帰分析の結果	46
表 4. 出生前変数と出生体重 SD スコアの関連－単変量回帰分析の結果	46
表 5. 出生前変数と在胎週数の関連－多変量回帰分析の結果	48
表 6. 出生前変数と出生体重 SD スコアの関連－多変量回帰分析の結果	48
表 7. 在胎週数関連 CpG における、KEGG パスウェイでのエンリッチメント 解析結果	54
表 8. FDR 基準で採択した在胎週数関連発現プローブにおける、KEGG パスウ エイでのエンリッチメント解析結果	61
表 9. [補正なし p 値] <0.05 の基準で採択した在胎週数関連発現プローブにおけ る、KEGG パスウェイでのエンリッチメント解析結果	62
表 10. 遺伝子発現解析に用いた臍帯血 55 サンプルの臨床データの分布	64
表 11. 在胎週数関連エピジェネティックメモリーの候補 CpG 2,093 箇所にお ける 25 種クロマチン状態のエンリッチメント解析結果	75
表 12. 在胎週数関連エピジェネティックメモリーの候補 CpG の中で発現との	

相関を認めた CpG63 箇所のクロマチン状態および DNA メチル化変化、遺伝子
発現との相関、生後のメチル化維持に関するまとめ・・・・・・・・・・ **81**

表 13. エピジェネティックメモリーの候補絞り込みの解析に臍帯血・生後児
血の DNA メチル化データを用いた 47 人の児の臨床データの分布・・・・・・・・ **83**

略語一覽

ART: Assisted Reproductive Technology

ASN_MAF: Asian Minor Allele Frequency

BMI: Body Mass Index

CAM: Chorioamnionitis

CI: Confidence Interval

Est.: Estimates

EWAS: Epigenome-Wide Association Study

FDR: False Discovery Rate

GA: Gestational Age

GDM: Gestational Diabetes Mellitus

IGR: Intergenic Region

iPROM: Idiopathic PROM

IVF: In Vitro Fertilization

kb: kilo base

kbp: kilo base pair

MAF: Minor-Allele Frequency

OR: Odds Ratio

PROM: premature rupture of the membrane

QC: Quality Control

RIN: RNA Integrity Number

SD: Standard Deviation

SD-score: Standard Deviation score

SGA: Small for Gestational Age

SNP: Single Nucleotide Polymorphysm

TF: Transcription Factor

TSS: Transcription Starting Site

UTR: Untranslated Region

要旨

出生体重 2,500g 未満の低出生体重児は我が国では 10%、OECD 加盟国全体で 6~7%の出生率を占める。低出生体重児は、在胎週数の早い早産児と出生体重 SD スコアの小さい Small for Gestational Age (SGA)児に分類されるが、いずれも周産期に子宮内外で栄養不良や低酸素の悪環境に曝露され、生存のための適応を余儀なくされることが多い。この「適応の過程で、成人後の疾病体質が形成される」のが DOHaD 仮説の概念であるが、エピジェネティックなメカニズムが大きく関与しているのではないかと考えられている。早産児・低出生体重児の分野では、DNA メチル化と臨床情報との関連を調べるエピゲノムワイド関連解析(EWAS)は世界中で行われており、知見も蓄積されているが、遺伝子発現との関連はほとんど検討されておらず、周産期に起こったメチル化変化の持続(エピジェネティックメモリー)に関しては結論が出ていない。本研究では、在胎 23 週~41 週の 110 人から得られた新生児の臍帯血・生後末梢血を用いて、単核球分離後サンプルについて EWAS および遺伝子発現解析を行い、早産児・SGA 児に起こるエピゲノム変化を検討し、エピジェネティックメモリーの候補をリストアップし、その特性について検討した。本研究は、早産児・低出生体重児の分野で血液サンプルを用いた我が国最初の EWAS 研究である。

I. 序文

1. 研究の背景

1-1. 低出生体重児の2大分類—早産児と Small for Gestational Age

(SGA) 児

出生体重が 2,500g 未満の低出生体重児の我が国の出生率は 10% に上り、OECD 加盟国全体でも 6~7% と一定の頻度での出生を認める。在胎週数と出生体重はともに低出生体重児の短期予後・長期予後を予測する上で最も重要な要素である。低出生体重児は大きく在胎週数が 37 週未満の早産児と、出生体重が在胎週数に比して 10 パーセンタイル未満 (およそ-1.28SD 未満に相当) の Small for Gestational Age (SGA) 児の 2 つのグループに分類される。早産児は妊娠 37 週以前に出生する児と定義されるが、本来胎児発育段階であるにも関わらず、正期産児に比べて早期に母体からの栄養や酸素を受け取ることができなくなり子宮外で生き抜くことを余儀なくされる。対して、SGA 児は子宮内で低酸素・栄養不良の状況に曝露されることで発生する傾向にある。これら 2 つのカテゴリの低出生体重児の間には病因や曝露の性質の違いはあるが、ともに周産期に酸素や栄養の供給が攪乱された状態にある。

1-2. 低出生体重児と **Developmental Origin of Health and Diseases**

(DOHaD)、そしてエピジェネティクスとの関連

これらの未熟児、ハイリスク児は後生に慢性疾患に罹患しやすいのだが、その疫学的な evidence はこれまでも蓄積している。例えば、早産児や SGA 児は冠動脈疾患、2 型糖尿病、メタボリックシンドローム、神経行動学的な問題を抱えることが多い[1-3]。オランダ飢饉に関する研究により、妊娠中の母体低栄養の後に低出生体重児が増加し、その低出生体重児は成人後の死亡率や疾病罹患率が高くなるという事実が補強された[4]。さらに、早産児および低出生体重児は小児期の後期でさえ、BMI が増加したり[5]、インスリン感受性が低下したり[6]、血圧が上昇する[7]など代謝動態に変化が生じる傾向にあることも分かっている。これらの事実は DOHaD 仮説(Developmental Origin of Health and Diseases)、すなわち胎児期・生後早期の悪環境を乗り越えるために起こる適応反応により成人期以降の慢性疾患への罹患しやすさが起こるという仮説を支持するものである[8, 9]。DOHaD 仮説は、「胎児期の低栄養・低酸素に晒された低出生体重児特に SGA 児が、子宮内プログラミングにより冠動脈疾患や 2 型糖尿病、メタボリックシンドロームなどの成人後の慢性疾患に罹患しやすくなる」とする Barker 仮説[2]を基盤として発展されたものであるが、「胎児期から幼少期の発育・発達

途中にかけて悪環境に晒された胎児・新生児・小児において、環境という外力により“くせ”ないし形質ひいては疾病体質が形成される」概念、“developmental plasticity”を基本概念とする仮説[9, 10]であり、対象とする疾病は上述のものに加え更にアレルギー疾患[11]、炎症性疾患[12]、精神疾患[13]、骨粗鬆症[9]、悪性腫瘍[14]など疾患全般が含まれる(表 1)。エピジェネティックなメカニズム、つまり DNA 塩基配列の変化によらない遺伝子発現の変化のメカニズムは細胞・組織の分化に中心的な役割を担う[15, 16]が、同時に環境曝露による“くせ”(形質・疾病体質)の形成の上でも重要な役割を果たし developmental plasticity の本質を説明するのに最も適したメカニズムの一つ[9, 10]ではないかと考えられている。すなわち、胎児期・生後早期におけるエピジェネティックな変化が後生の慢性疾患の発症に関与しているかもしれない。

表 1. DOHaD 仮説と Barker 仮説の関係

	Barker仮説	DOHaD仮説
基本概念	子宮内プログラミング	Developmental Plasticity
対象	低出生体重児、主にSGA児	環境曝露を受ける全ての胎児・新生児・小児
曝露の期間	胎児期	胎児期～幼児期
成人後疾病	虚血性心疾患、2型糖尿病、肥満など主に生活習慣病	疾患全般: 左の疾患に加え、アレルギー疾患、骨粗鬆症、精神疾患、炎症性疾患、悪性腫瘍など

1-3. これまでの早産児・低出生体重児領域のエピゲノムワイド関連研究

前述の通りエピジェネティックなメカニズムは developmental plasticity の上で重要な役割を果たすと考えられている。エピジェネティクスには DNA メチル化、ヒストン修飾、micro RNA、クロマチンリモデリングなどが含まれ、いずれのメカニズムも重要である[9, 17]が、このうち最も実験系として安定してゲノムワイドな情報を供給でき、分子疫学的検討が行いやすいのが、DNA メチル化情報[18]である。また、ChromHMM[19, 20]に代表にされるクロマチン修飾などの推定ソフトウェアなどにより他のエピジェネティックな修飾の推定も DNA メチル化解析結果から可能になってきている。EWAS(Epigenome-Wide Association Study: エピゲノムワイド関連分析)とは、現状では DNA メチル化と臨床情報の関連をゲノムワイドに分子疫学的に検討することを指すのがほとんど[18]だが、マイクロアレイ技術の進歩により世界中で精力的に行われている。周産期の分野では臍帯血を用いた研究で、在胎週数[21-26]、出生体重[23, 27]、出生体重 z-score[28]に関連するメチル化変化を検討した報告がこれまでにされている。しかしながら、これらの研究では、2点明らかになっていないことがある。第一に、DNA メチル化変化が RNA 発現に関与するかどうかに関しては検

討されていない。典型的には転写開始点の最近傍に位置し CpG サイトが豊富なことの多い Promoter 領域での DNA メチル化は遺伝子発現の抑制に働くとされ [9]、Gene Body 領域における DNA メチル化は遺伝子発現の増加に働くことが多い [29]とされているが、DNA メチル化と臨床情報の関連についての検討で終始していることがほとんどで、DNA メチル化・遺伝子発現の同時の検討はまだ行われておらず、実際の DNA メチル化による遺伝子発現への影響は明らかにされていない。第二に、生後のメチル化についての検討はほとんど行われておらず、検討されていたとしても研究によって解釈が異なる。先行研究 2 報によると出生時に確認された沢山のメチル化変化は成人期までには消失する [23, 30] が、これら 2 つの生後のメチル化変化の結果は異なり、早産および SGA に関連してエピジェネティックメモリーが生じるかどうかはコンセンサスの得られていないところである。その他、これらの先行研究の問題点として、検討した細胞は積極的に分離されておらず、全血サンプルないし血液中の全白血球および全有核細胞からなる細胞集団を意図したサンプル(以下、これらを「全血ないし全血に準じたサンプル」と呼ぶ)のメチル化を解析したものであることがあげられる。

2. 研究の目的

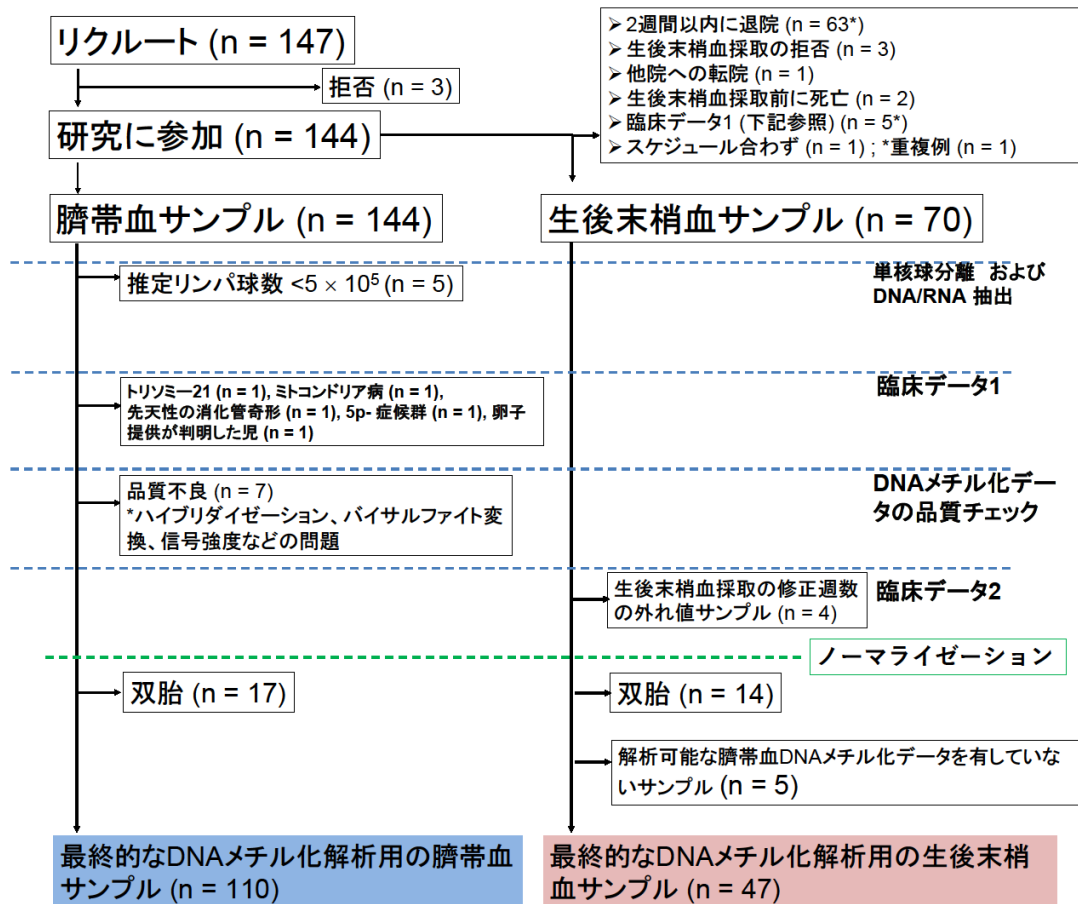
本研究の目的は DNA メチル化アレイおよび遺伝子発現アレイを用いて早産および SGA に関連したエピジェネティック変化を調べ、生後にも持続する変化すなわちエピジェネティックメモリーを探索することである。また、本研究ではより正確にエピジェネティック変化を解析するために、単核球分離後の buffy coat を用いることで可能な限り骨髓球を除外して細胞集団のリンパ球純度を上げ、DNA メチル化、遺伝子発現、およびそれらの関連を同時に評価した。本研究は早産および SGA 児の分野において、(1)日本における血液検体でのエピゲノムワイド関連解析(EWAS)を行った最初の研究であり、(2)同一の単核球分離後サンプル内で遺伝子発現も広範囲に検討した世界で初めての EWAS 研究である。

II. 方法

1. 研究対象の集団

本研究は前向き横断型観察研究である。計画当時(2014年6月)の先行研究([21, 22, 26, 30]; これらの先行研究のサンプル数の中央値は95.5)を基にして、在胎週数に関連する CpG が確認されることを期待して、臍帯血サンプル数が100以上で、早産児と正期産児を同程度の割合で含むことを目標にした。この目標の上、著者、研究協力者および参加者の母児の準備が整っている時間に該当する児を対象にリクルートした。2014年10月から2016年7月までの期間で東京大学医学部附属病院・東京都立墨東病院で出産した母児147組をリクルートし、144組から書面での同意を得た。(拒否は3組だった。) 図1(次頁)に研究参加者のリクルート、およびDNAメチル化解析サンプルのフィルタリングまでの流れを示す。全ての参加者から臍帯血を得た。臍帯血は児の娩出の際に採取した。対して、生後の児の末梢血に関しては修正36週(在胎週数に出生後週数を加算したときに36週となる時点)を過ぎて、採取が可能な場合限り、サンプリングを行った。出生後2週間以後に、入院を継続しており、呼吸・循環・貧血などを含む全身状態が安定していると担当医が判断し、最も予定日に近い日に静脈穿刺により生後末梢血のサンプリングを行った。

図 1. 研究参加者のリクルート、および DNA メチル化解析用サンプルのフィルタリングまでの流れ



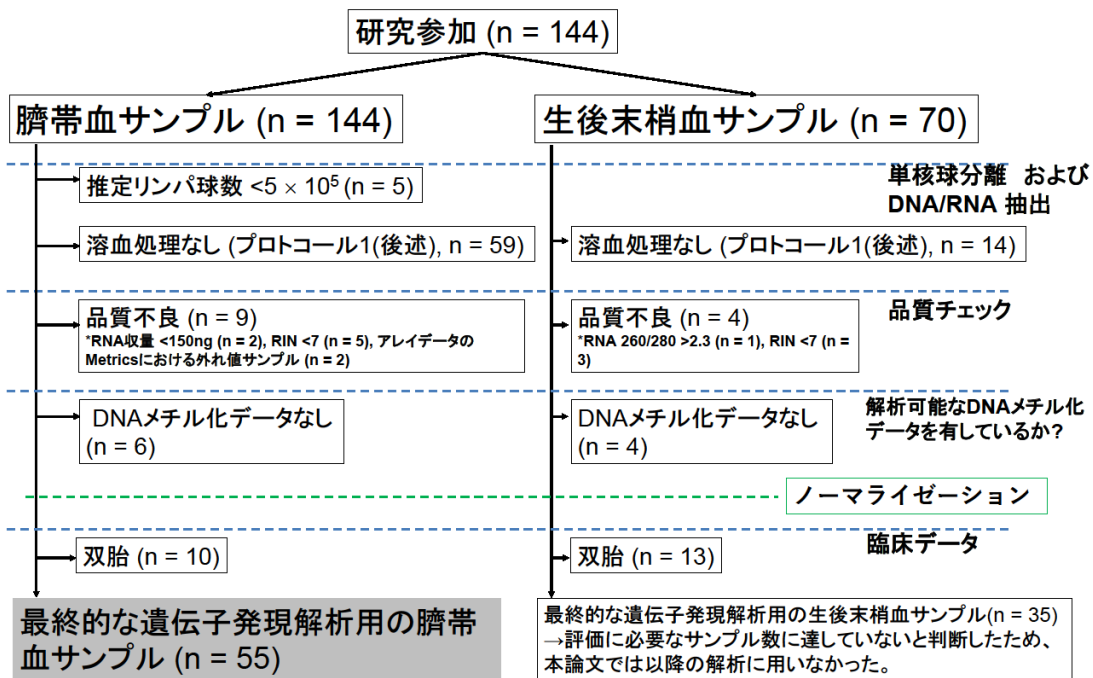
生後児血は以下の参加者からは得られなかった。； 出生後 2 週間以内に退院した児 (n = 63)、生後末梢血のサンプリングを両親が拒否した児 (n = 3)、他院に転院した児 (n = 1)、生後末梢血採取前に死亡した児 (n = 2)、出生後に先天性疾患があることが判明した児 (n = 4)、サンプリングの予定や都合がつけられなかった児 (n = 1)。得られた生後児血検体は 70 サンプルであった。DNA メチル化アレイ解析用のサンプルフィルタリングとして、得られた臍帯血 144 サンプルのうち、推定リンパ球数が 5×10^5 未満のもの (n = 5)、品質不良であったもの

(n = 7)、出生後に trisomy21 (n = 1)、ミトコンドリア病 (n = 1)、先天性消化管奇形 (n = 1)、5p-症候群 (n = 1)、卵子提供 (n = 1)が判明したものは除外し、また最終的にノーマライゼーション(統計解析を円滑にするために、解析前に実験誤差などのバイアスやノイズを取り除く観測値補正のステップ)の後に、遺伝的・環境的要因により解析結果を増強・減弱させるリスクのある双子のサンプル (n = 17) もエピゲノムワイド関連解析(EWAS)の推奨案[18]に従い最終的に除外した。生後児血 70 サンプルからはサンプリング時の修正週数が外れ値であった 4 サンプルを除外し、臍帯血と同様にノーマライゼーションの後に双子のサンプル (n = 14)および解析可能な臍帯血メチル化データを有していないサンプル(n=5)も除外した。最終的な解析用の DNA サンプルは臍帯血が 110 サンプル、生後児血が 47 サンプルであった。

次に遺伝子発現アレイ解析用の RNA サンプルのフィルタリングの流れについて述べる(図 2; 次頁)。臍帯血発現アレイ解析用のサンプルフィルタリングでは、得られた 144 サンプルのうち、推定リンパ球数(自動血球カウンターで総白血球数にリンパ球分画を乗じて算出)が 5×10^5 未満のもの (n = 5)、単核球分離の際に溶血剤処理を行わなかったもの (n = 59)、品質不良のもの (n = 9)、解析可能な DNA メチル化データを有していないサンプル (n = 6)を除外し、DNA

メチル化解析と同様にノーマライゼーションの後に双胎のサンプル (n = 10) も除外した。最終的な解析用の臍帯血 RNA サンプルは 55 サンプルであった。

図 2. 遺伝子発現アレイ解析用の RNA サンプルのフィルタリングの流れ



最後に、生後末梢血発現アレイ解析用のサンプルフィルタリングでは、得られた生後児血 70 サンプルのうち、単核球分離の際に溶血処理を行わなかったもの (n = 14)、品質不良のもの (n = 4)、解析可能な DNA メチル化データを有していないサンプル (n = 4) を除外し、DNA メチル化と同様にノーマライゼーションの後に双胎のサンプル (n = 13) も除外した。最終的な解析用の生後児血 RNA サンプルは 35 サンプルであり、評価に必要なサンプル数に達していないと判断したため、本論文では以降の解析に用いなかった。

2. 研究対象集団の臨床データの収集および算出

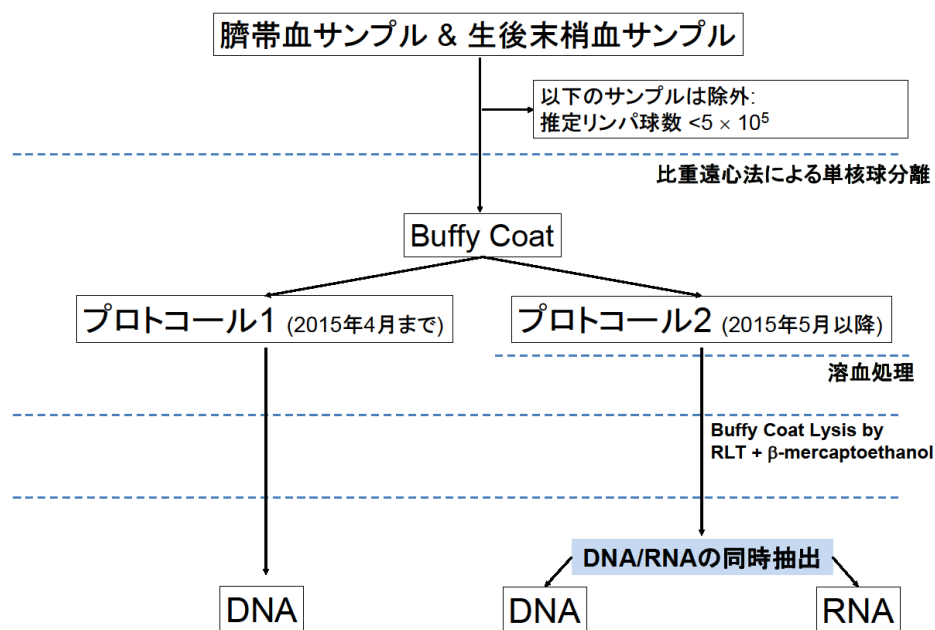
臨床データ(在胎週数(GA)、出生体重、児の性別、経産数、分娩様式(帝王切開 or 経膈分娩)、母体年齢、妊娠前母体 BMI、IVF(In Vitro Fertilization)など ART(生殖補助医療)による妊娠か否か、母体の喫煙状況(妊娠中母体喫煙、妊娠前母体喫煙)、母体の妊娠中の情報(妊娠糖尿病、絨毛膜羊膜炎、妊娠高血圧症、特発性の前期破水、前置胎盤の有無)、生後児血サンプリング時の修正週数)は病院診療録から得た。父親のデータ(体重、身長、年齢)は質問紙を通して収集した。出生体重の SD-score、すなわち日本のレファレンスデータに基づいた在胎週数に比しての出生体重の z-score は、日本小児内分泌学会が提供している計算プログラム(<http://jspe.umin.jp/medical/keisan.html>)により、在胎週数、出生体重、児の性別、経産数の入力から算出した。

3. 単核球分離と DNA および RNA の抽出

採取した臍帯血および生後末梢血から採取後 24 時間以内に Ficoll 比重液を用いた比重遠心法による単核球分離を行った。(図 3; 次頁) 血液サンプルの容量が 1ml 未満のときには、自動血液分析装置を用いて白血球数と全白血球中のリンパ球分画を測定した。推定のリンパ球数が 5×10^5 未満のサンプルは単核球

分離の処理は行わなかった。

図 3. 単核球分離および DNA/RNA 抽出までの工程



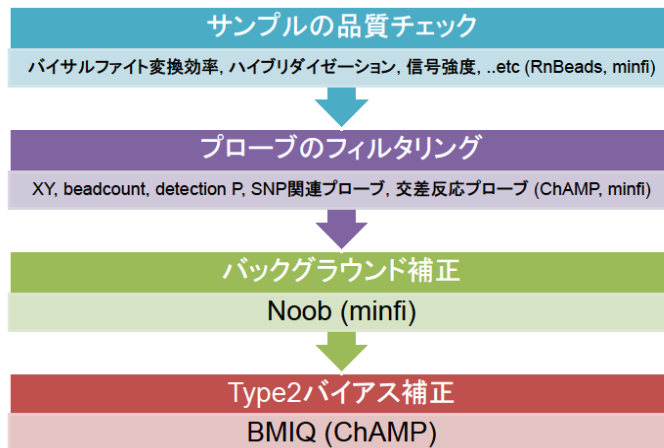
2015年4月までに得た59個の臍帯血サンプルと14個の生後児血サンプルに関しては、遠心後のbuffy coatを直接 β -メルカプトエタノール添加のRLT bufferでlysisした。ゲノムDNAは単核球分離後のbuffy coatのライセートからQiagen社のAllPrep kitを用いて抽出した(プロトコール1)。2015年5月以降に得た残りの臍帯血76サンプルと生後児血52サンプルに関してはプロトコール2を適用した。プロトコール2では最初に遠心後のbuffy coatを赤血球溶血液に混ぜた。2番目にこの混ぜた溶血液を遠心してペレットから赤血球を含む溶血液を除去した。3番目にペレットを β -メルカプトエタノール添加のRLT bufferでlysisした。4番目にペレットのライセートから、プロトコール1で用いたのと同じキッ

トを用いてゲノム DNA とトータル RNA を同時抽出した。最後に、プロトコール 1 とプロトコール 2 とともに、ゲノム DNA とトータル RNA のアリコートをマイクロアレイ実験まで -80°C で保存した。

4. DNA メチル化アレイにおける Quality Control とデータの前処理

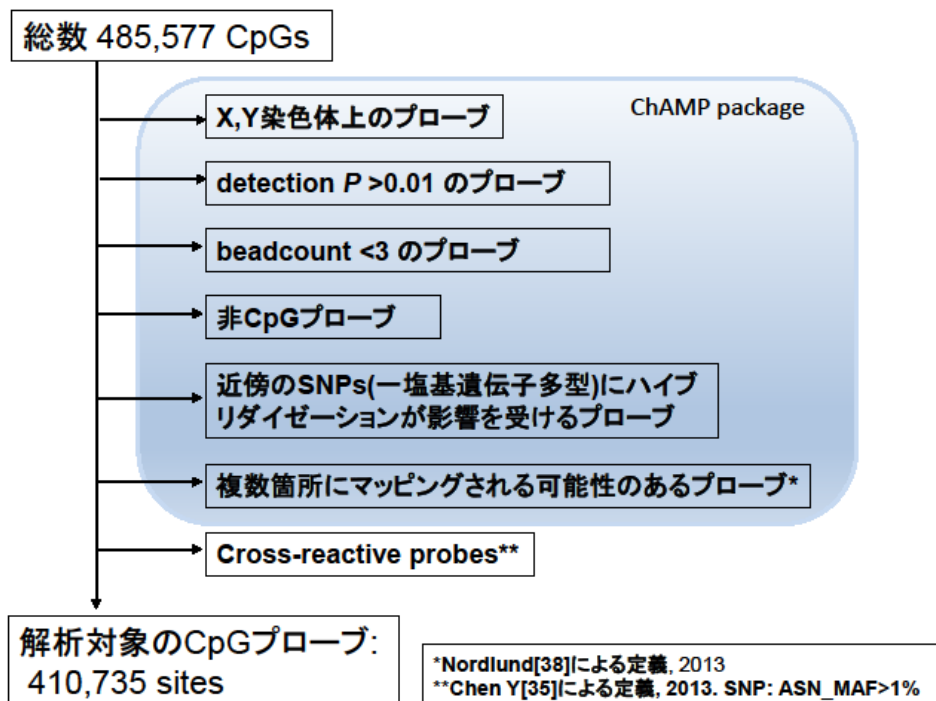
抽出の後、ゲノム DNA を EpiTect Plus DNA Bisulfite kit (Qiagen 社)を用いてバイサルファイト変換した。このバイサルファイト DNA を、48 万カ所以上の CpG サイトのメチル化レベルを調べることのできる DNA メチル化アレイ、Infinium HumanMethylation450 BeadChip (イルミナ社)[31]を用いてデータ処理した。各 CpG サイトのメチル化レベルは β 値として算出した。 $(\beta = \text{メチル化アレルの信号強度 } (M) / (\text{非メチル化アレルの信号強度 } (U) + \text{メチル化アレルの信号強度 } (M) + 100))$ すなわち、 β 値は 0 (完全な非メチル化)から 1 (完全なメチル化)までの範囲の値をとることになる。全てのメチル化データの前処理は R 環境 (version 3.3.2)内で行なった。各サンプルの品質は RnBeads package(v. 1.4.0)[32]および minfi package(v. 1.20.2)[33]を用いて行い、バイサルファイト変換効率、ハイブリダイゼーション効率、メチル化プローブおよび非メチル化プローブの信号強度の基準に基づき、品質不良と判断したプローブは除外した(図 4)。

図 4. DNA メチル化アレイにおける品質チェックとデータの前処理



Quality Control (QC: 品質チェック)に合格した DNA メチル化解析用のデータセットは 127 個の臍帯血サンプルと 66 個の生後末梢血サンプルで構成されている。品質不良のサンプルを除いた後、ChAMP package(v.2.6.0)[34]を用いて probe filtering を行なった(図 5)。

図 5. DNA メチル化解析における Probe Filtering



品質チェック用に搭載されている非 CpG プローブは解析対象から除外した。XY 染色体上にも重要な遺伝子が含まれているが、男女共通に起こる DNA メチル化変化を検討する研究デザインであるため、また交差反応を起こしやすく評価の際の注意が必要になることが多い[35]ため、先行研究[23, 25, 28, 30, 36, 37]に倣って XY 染色体上のプローブは除外した。その他、**detection p -value**(≡アレイ画像内でのプローブシグナルとバックグラウンドの識別度)が 0.01 以上のプローブ、ビーズカウントが 3 未満のプローブは除外した。ChAMP package の初期設定に基づき、Nordlund[38]が定義した複数箇所にマッピングされる可能性のあるプローブは除外した。さらに、Chen ら[35]のデータに基づき、交差反応する恐れのあるプローブやアジア人の MAF(minor-allele frequency)が 1%以上の多型関連プローブは除外した。Probe filtering の後、410,735 個の CpG プローブが以降の解析対象として残った。バックグラウンド補正および dye-bias(染め上げバイアス)補正 (Noob)[39]を minfi package を用いて行った後、type2 probe のバイアスを減らすために beta-mixture quantile normalization (BMIQ)[40]を ChAMP package を用いて行った。これら全ての前処理の後、最終的に双胎のサンプル(臍帯血 17, 生後末梢血 14)をエピゲノムワイド関連解析(EWAS)の推奨案[18]に従い除外した。

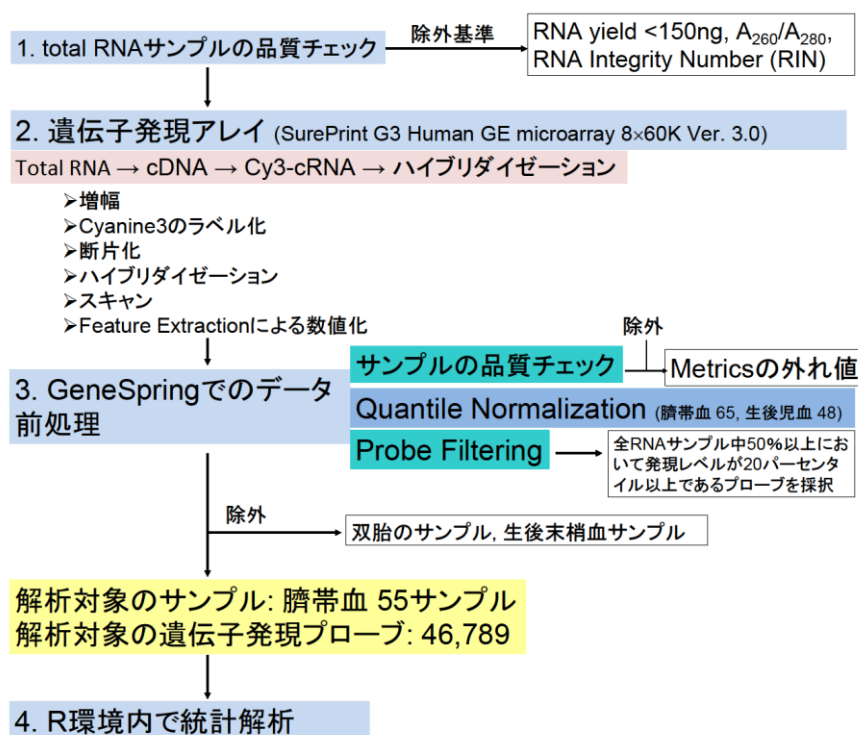
5. アレイデータによる単核球分離後 buffy coat 内の細胞分画の推定

単核球分離後 buffy coat 内の細胞分画(白血球と有核赤血球の組成)について臍帯血解析用の Bakulski[41]のレファレンスデータセットを用いて Houseman[42]が提案した方法を用いて推定値を算出した。リンパ球(CD4+T 細胞、CD8+T 細胞、NK 細胞、B 細胞)、骨髄球(顆粒球、単球)、および有核赤血球(赤芽球)を含む細胞分画の推定値データは以降の多変量メチル化解析に用いた。

6. 遺伝子発現アレイとデータ前処理

遺伝子発現アレイとデータ前処理の流れを図6に示す。

図6. 遺伝子発現アレイ解析とデータ前処理の流れ



抽出の後に、最初にトータル RNA を NanoDrop-1000 spectrometry (Thermo Scientific) を用いて、定量および品質チェックを行なった。RNA 収量が 150ng 未満のものや、RNA 吸光度比(A_{260}/A_{280})が 2.3 以上のサンプルは除外した。トータル RNA の品質は更に Agilent 2100 Bioanalyzer system (Agilent Technologies) の自動キャピラリー電気泳動を用いて確認した。RNA integrity number (RIN)が 7 未満のサンプルは除外した。そして、トータル RNA100ng をもとに Low Input Quick Amp Labeling Kit (Agilent)を用いて Cyanine 3 でラベル化した cRNA を合成した。ラベル化の後、cRNA 600ng を Agilent Gene Expression Hybridization Kit (Agilent) を用いて断片化した。断片化した cRNA を SurePrint G3 Human GE microarray 8×60K Ver. 3.0 (Agilent Technologies)にハイブリダイズした。ハイブリダイゼーションの後、Gene Expression Wash Buffer の 1 および 2 (Agilent Technologies)を用いて、アレイスライドをウォッシュし、Agilent DNA Microarray Scanner (Agilent Technologies)でスキャンした。Agilent Feature Extraction (FE) software (ver. 10.7.3.1) を用いて、raw intensity data を得た。

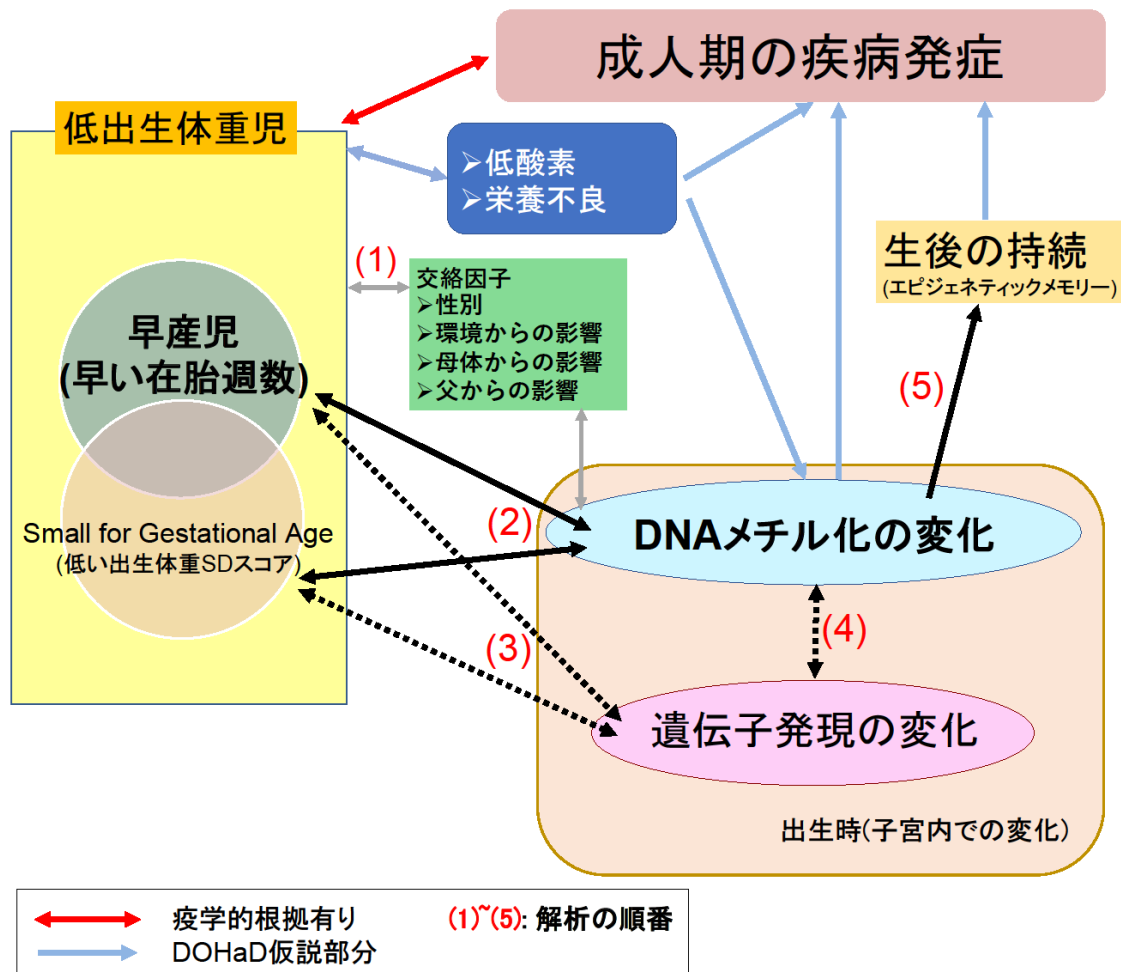
各サンプルのアレイの品質チェック、probe filtering 、ノーマライゼーションは GeneSpring14.5 (Agilent Technologies)を用いて行なった。Metrics で外れ値を示したサンプルは除外した。遺伝子発現解析のためのサンプル QC を通過したデータ

セットは臍帯血 65 サンプル、生後末梢血 48 サンプルで構成した。サンプル QC の後に、このデータセットで `quantile normalization` を行なった。全 RNA サンプル中 50% 以上において発現レベルが 20 パーセントイル以上であるプローブを採択し、XY 染色体上のプローブは除外した。46,789 プローブが以降の解析に使用できることとなった。これら全ての前処理の後、最終的に EWAS の推奨案[18] に準じて双胎のサンプル(臍帯血 10, 生後児血 13)を除外した。生後児血の RNA サンプルは 35 サンプルであり、評価に必要なサンプル数に達していないと判断したため、本論文では以降の解析に使用しなかった。ノーマライゼーションとフィルタリングの工程を経たデータを R 環境に取り込み、統計解析した。

7. 統計解析

全ての統計解析は R(v. 3.3.2)環境内で行なった。全体の解析の枠組みは図 7(次頁)にまとめ、各解析の詳細については、以下 7-1 ~ 5 に記載した。

図 7. 全体の解析の枠組み



7-1. 在胎週数ないし出生体重 SD スコアと関連のある臨床的特徴の

検討

最初に、児の在胎週数(GA)または出生体重 SD スコアと関連のある出生前の変数(臨床的特徴)を調べるために単変量回帰分析を行なった。在胎週数、出生体重 SD スコア、児の性別に加え、在胎週数または出生体重 SD スコアに関連する 6 つの出生前の変数を以降の解析に用いることとした。次に、在胎週数また

は SD スコアとの関連を再確認するために、目的変数を除いた全ての変数で調整した多変量回帰分析を行なった。Gelman[43]と Lin ら[37]が用いた方法に従って、二値出生前変数の正規化は行わなかった。また連続出生前変数は標準偏差が 0.5 となるように標準化して、連続変数と二値変数それぞれの効果量を比較できるようにした。

7-2. 在胎週数ないし出生体重 SD スコアについてのエピゲノムワイド関連解析(EWAS) およびパスウェイ解析

全 CpG プローブにおける説明変数と臍帯血の DNA メチル化変化を評価するために、2つの多変量線形回帰モデルによる感度分析的手法を用いた解析を行った。説明変数は在胎週数(GA)および出生体重 SD スコアである。

$$Y_{meth} = \beta_0 + \beta_{GA} \times X_{GA} + \beta_{SDscore} \times X_{SDscore} + \sum (\beta_i \times X_{cov}) + \varepsilon$$

2つの線形回帰モデルのうち、一つのモデルは各 CpG サイトの DNA メチル化変化を目的変数、児の性別・バッチ・推定細胞分画など最低限必須と考えられる項目を調整変数として検討するモデルである。対して、もう一つのモデルは前者のモデルに絨毛膜羊膜炎・特発性の前期破水・妊娠高血圧症・母体喫煙状況(母体の妊娠前喫煙)・母体の妊娠前 BMI・帝王切開を調整変数に加えたもので

ある。前者のモデルが主要解析モデルの「モデル 1」、後者がサブ解析モデルの「モデル 2」である。サンプル数が 110 と限られているため、調整変数の項目数としてはモデル 1 の方が適切であるが、可能な限り交絡調整を行う目的で感度分析の概念に基づきモデル 2 での確認を行なった。具体的な手順としては、まず各解析モデルにおいて、410,735 箇所の CpG サイトの多重検定補正を行い、false discovery rate (FDR)が 0.05 未満[44]であるものを在胎週数ないし出生体重 SD スコアに関連する CpG サイトとしてまず採択した。そして最終的にモデル 1 とモデル 2 で両方とも在胎週数ないし出生体重 SD スコアと関連を認め、モデル 1 とモデル 2 での回帰係数の符号が一致するものを在胎週数関連 CpG サイトないし SD スコア関連 CpG サイトとして決定した。

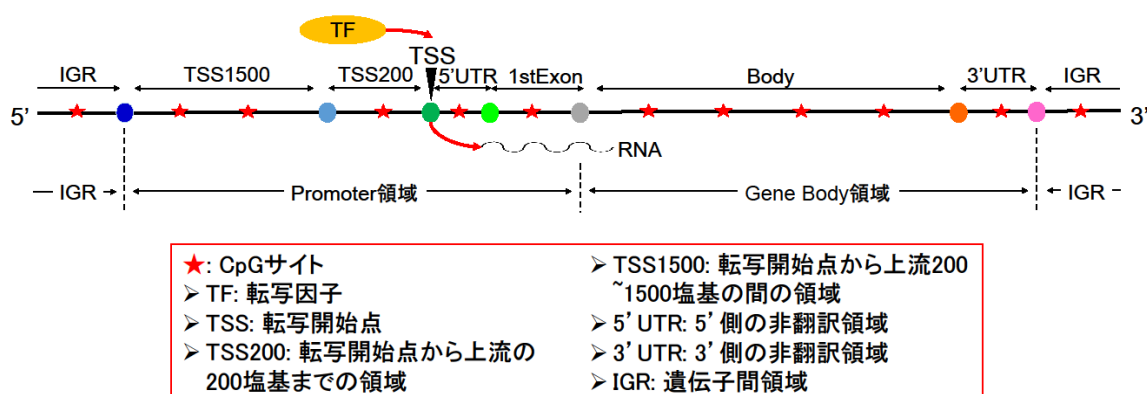
エピゲノムワイド関連解析(EWAS)の後、パスウェイ解析を行った。在胎週数関連 CpG サイトないし SD スコア関連 CpG サイトと遺伝子の紐付けは ChAMP package [34]を用いて行った。次に回帰係数の符号に従って、在胎週数関連 CpG サイトをメチル化率増加型およびメチル化率減少型在胎週数関連 CpG サイトに、SD スコア関連 CpG サイトをメチル化率増加型およびメチル化率減少型 SD スコア関連 CpG サイトに分類した。イルミナ社の遺伝子のアノテーションがないプロンプや gene body, 3'UTR(3'側の非翻訳領域), intergenic region (IGR: 遺伝子間領

域)にマッピングされるものはパスウェイ解析には用いなかった。また、重複する遺伝子名は一つと数えた。各カテゴリーにおいて遺伝子名が 2,000 以上になった場合は遺伝子上に 2 つ以上の関連 CpG サイトがある遺伝子名のみをパスウェイ解析に用いた。最後に DAVID Bioinformatics Resources 6.8 [45] で KEGG パスウェイにおけるエンリッチメントを評価した。多重検定補正後 FDR の基準に基づいた評価を行うために Benjamini-Hochberg 法を用いた。パスウェイ解析での採択基準は enrichment-FDR 0.1 以下とした。

在胎週数関連 CpG ないし出生体重 SD スコア関連 CpG の所在について、転写開始点を軸としたゲノム領域分類、CpG アイランドを軸とした分類で整理したので、これらの分類とその一般的な意義についてもこの項で説明する。まず、前者は遺伝子の転写開始点からの位置関係、つまり遺伝子の構造をベースにした CpG サイトの所在の分類である(図 8; 次頁)。遺伝子は大きく Promoter 領域 (TSS1500, TSS200, 5'UTR, 1stExon)、Gene Body 領域(Body, 3'UTR)に別れ、その他は遺伝子間領域(IGR: Intergenic Region)となる。定説では Promoter 領域の DNA メチル化は遺伝子発現の低下につながる[9]とされ、また Gene Body 領域の DNA メチル化は遺伝子発現の亢進につながる人が多い[29]と言われている。この分類での DNA メチル化の遺伝子発現への寄与の度合いを大小関係で表すと、

[Promoter 領域] > [Gene Body 領域] > [遺伝子間領域]、となる。

図 8. 転写開始点(TSS)を軸とした CpG サイトの所在に関するゲノム領域分類



次に、CpG アイランドを中心とした分類について説明する(図 9)。CpG アイラン

ドは 200 塩基以上の長さを有し、GC 部分が全体の 50%以上を占める 領域で、

CpG shore は CpG アイランドの一端から 2,000 塩基の間領域、CpG shelf は CpG

アイランドの一端から 2000~4000 塩基の領域、Open Sea は左のどのカテゴリ

ーにも属さず、GC 部分の少ない領域になる。前述の Promoter 領域の 7 割が CpG

アイランドを有するとされる[46]が、この CpG アイランドは低メチル化状態に

あることが多い[47]のだが、実際にはこの CpG アイランドのメチル化変化は起

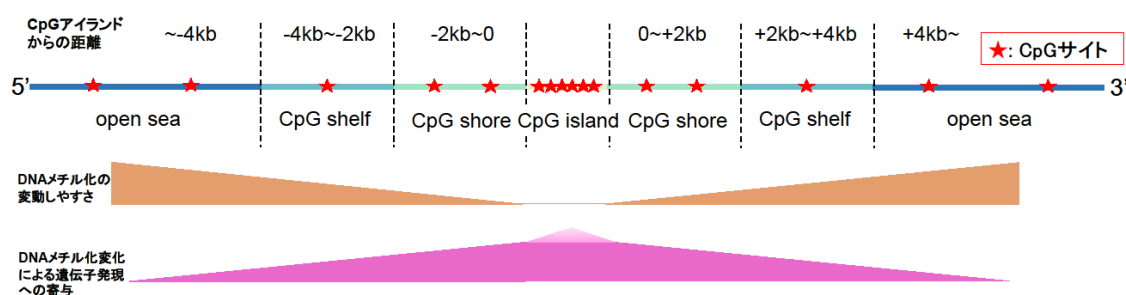
こりにくく、むしろ CpG アイランドの外側の領域の方が、変化が起こりやすい

のだが、これら外側のメチル化変化も遺伝子発現の調節に関与する[48]と言われ

ている。しかし、外側の変化も遺伝子発現に関係するとはいえ、やはり CpG ア

イランドからの距離が近い方が遺伝子発現への影響は一般的には強い。

図 9. CpG アイランドを軸とした CpG サイトの所在に関する分類



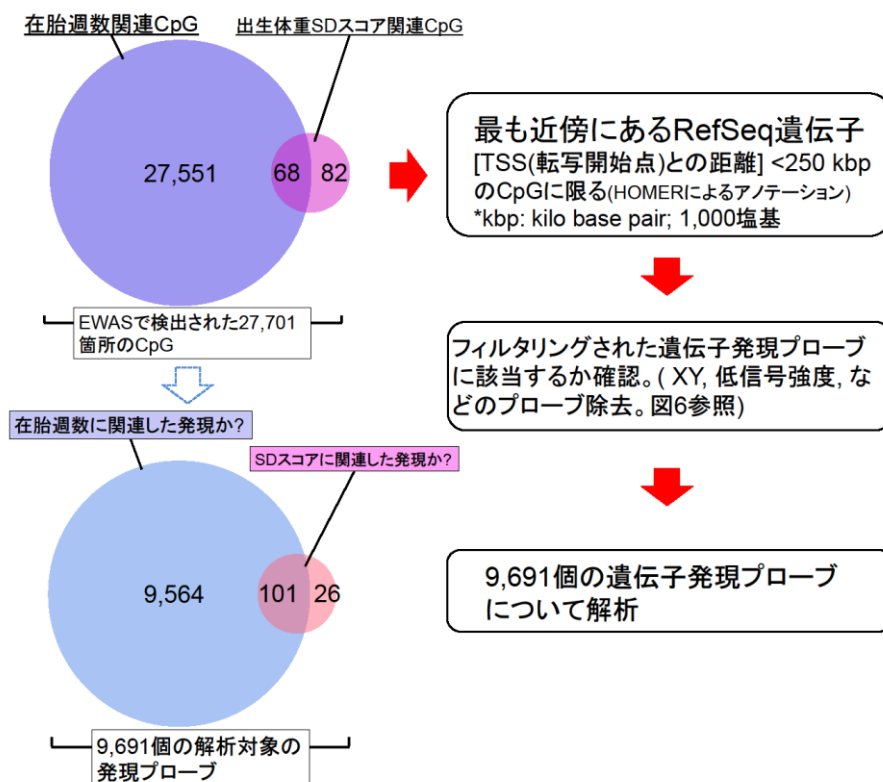
7-3. 在胎週数ないし出生体重 SD スコアについての遺伝子発現関連解析お

よびパスウェイ解析

在胎週数ないし出生体重 SD スコアに関連する遺伝子発現の変化を探索するために、臍帯血の EWAS でメチル化変化を認めた遺伝子を対象に遺伝子発現アレイ解析を行った。遺伝子発現関連解析の準備にあたって、CpG サイトを対応する遺伝子発現プローブにマッチングさせる必要があったが、Promoter 領域だけでなく Gene Body 領域も発現に影響を及ぼしうること[29, 49-51]、遺伝子発現の調節に強く影響すると考えられている CpG アイランドのメチル化変化は実際には起こりづらくその周囲の DNA メチル化変化が頻度的には重要であり遺伝子発現の調節にも関わること[48, 52, 53]、転写開始点あるいは CpG アイランドからのどの程度離れば遺伝子発現との関連がなくなるかという明確ある

いは普遍的な線引きはまだできていないことから、他の分野の先行研究[54, 55]に倣って、同一染色体上での塩基同士の相互作用(*cis*-interaction)がおよそ 250kb 上下流の範囲で起こっていた知見[56]を考慮し、可能な限り全ての CpG サイトを 250kb 上下流の範囲内に位置する品質チェック済みの遺伝子発現プローブにマッチさせることとした。こうして、在胎週数ないし出生体重 SD スコアと FDR 基準で関連を認めた 27,701 箇所の CpG プローブの中で、Homer のアノテーションツール[57]を用いて、19,790 プローブを対応する遺伝子発現プローブにマッチさせた(図 10)。

図 10. EWAS で得られた在胎週数関連 CpG と出生体重 SD スコア関連 CpG を、対応する遺伝子発現プローブにマッチング



その後、臍帯血での EWAS で使用した「モデル 1」と同様の多変量線形回帰モデルを用いて、説明変数と臍帯血における遺伝子発現強度(底 2 の対数変換後のもの)の関係を評価した。(「モデル 2」による解析は、調整変数の数がサンプル数 55 に比して多すぎて不適切と判断したため、遺伝子発現関連解析には用いなかった。) 説明変数は在胎週数(GA)および出生体重 SD スコアである。9,691 個の遺伝子発現プローブの多重検定補正を行い、false discovery rate (FDR)が 0.05 未満[44]であるものを在胎週数ないし出生体重 SD スコアに関連する発現プローブとして採択した。

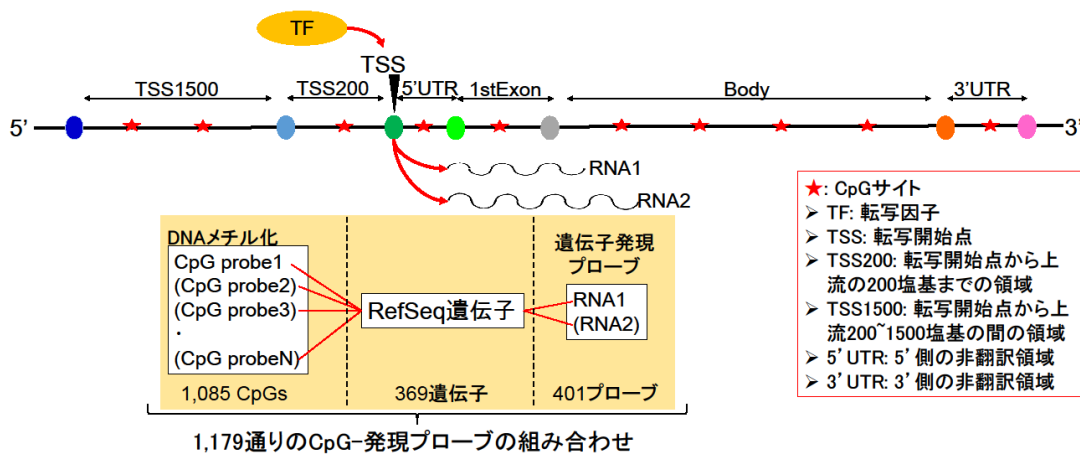
多変量遺伝子発現解析の後、FDR 基準(FDR <0.05)で関連のあったプローブ群を対象とするパスウェイ解析と、FDR 基準での説明変数との関連プローブが存在した場合に限り追加で、多重検定補正なしで関連があったプローブ群を対象にパスウェイ解析を行なった。本研究では回帰係数の符号に従って、在胎週数関連発現プローブを発現増加型および発現減少型在胎週数関連発現プローブに、SD スコア関連発現プローブを発現増加型および発現減少型 SD スコア関連発現プローブに分類した。アジレント社の遺伝子のアノテーションがないプローブはパスウェイ解析には用いなかった。また、重複する遺伝子名は一つと数えた。最後に DAVID Bioinformatics Resources 6.8 [45] で KEGG パスウェイに

おけるエンリッチメントを評価した。多重検定補正後 FDR の基準に基づいた評価を行うために Benjamini-Hochberg 法を用いた。パスウェイ解析での採択基準は enrichment-FDR 0.1 以下とした。

7-4. DNA メチル化と遺伝子発現の関連の確認

DNA メチル化と遺伝子発現の直接の関連を確認するために、本研究では FDR 基準(FDR <0.05)で在胎週数と関連のある CpG サイト(臍帯血での EWAS)および FDR 基準で在胎週数と関連のある遺伝子発現(臍帯血での遺伝子発現関連解析)をともに有する遺伝子について DNA メチル化と遺伝子発現の相関をみる解析を行った。最初に、Homer のアノテーションツール[57]を用いて、369 個の RefSeq 遺伝子に由来する 401 個の在胎週数関連転写産物、および対応する 1085 箇所の CpG で構成される 1,179 個の CpG と遺伝子発現の組み合わせを、遺伝子発現関連解析の準備と同様に、250kb 上下流の範囲内であれば転写開始点や CpG アイランドなどを軸とした領域区分による制限は特に行わず、総当たりで作成した(図 11; 次頁)。次に、1,179 個の各組み合わせに対して単変量線形回帰モデルを用いて関連性を調べた。ここでは、補正なしの p 値においての 0.05 未満を有意基準とした。

図 11. DNA メチル化-遺伝子発現の直接比較のために、在胎週数関連 CpG サイトと 対応する在胎週数関連発現プローブの組み合わせを作成 (n=55, 臍帯血)



7-5. エピジェネティックメモリーの候補絞り込み

在胎週数ないし出生体重 SD スコアについての EWAS の結果をもとに、本研究では出生時にみられたメチル化変化が生後も持続(エピジェネティックメモリー)している CpG サイトの候補を選ぶ為の解析を行った。臍帯血と生後末梢血の両方のデータを得られる 47 人の児のサンプルを用いて、在胎週数関連および出生体重 SD スコア関連 CpG サイトに対して、臍帯血サンプルと生後末梢血の間メチル化率の Pearson の相関係数を算出した。本研究では相関係数が 0.7 以上の CpG サイトをエピジェネティックメモリーの候補と考えた。そして候補 CpG サイトが 50 箇所以上であった場合には、ChromHMM[20]において 12 種類のヒストンマークで特徴づけられる 25 種類のクロマチン状態のエンリッチメン

トを調べる解析を行った。このエンリッチメント解析では、本研究では臍帯血 T 細胞に基づいたアノテーションと臍帯血 B 細胞に基づいたアノテーションを利用した。この 25 種類のクロマチン状態のエンリッチメントの評価は Fisher の正確検定を用いて行い、多重検定補正には Bonferroni 法を用いた。Bonferroni 補正後 p 値(ここでは[補正前 p 値] × 25)の有意基準は 0.05 未満とした。

Ⅲ. 結果

我々は臍帯血 110 サンプルと生後末梢血 47 サンプルからのノーマライズされた DNA メチル化データ、および臍帯血 55 サンプルからのノーマライズされた遺伝子発現データを得た。全ての解析用データは単胎の児由来である。これらの解析に参加していただいた 110 組の母児において、平均の在胎週数は 34 週で平均の SD スコアは-0.6 であった(表 2)。

表 2. 本研究における臍帯血 110 サンプルの臨床データの分布

	平均値 (SD)	中央値	N (%)
在胎週数	34.0 (4.9)	35.0	
<28 週			17 (15.5)
28-32 週			17 (15.5)
32-37 週			31 (28.2)
>37 週			45 (40.9)
出生体重 SD スコア	-0.6 (1.4)	-0.5	
<-2.5			16 (14.5)
-2.5 ~ -1.28			22 (20.0)
-1.28 ~ 1.28			60 (54.5)
1.28 ~ 2.5			12 (10.9)
性別			
男児			53 (48.2)
女児			57 (51.8)
分娩様式			
帝王切開			89 (80.9)
経膣			21 (19.1)

	平均値 (SD)	中央値	N (%)
母体妊娠前 BMI	21.1 (3.5)	20.3	
<18.5 kg/m ²			23 (20.9)
18.5 ~ 25 kg/m ²			77 (70.0)
25 ~ 30 kg/m ²			4 (3.6)
>30 kg/m ²			6 (5.5)
妊娠中の母体喫煙			
Yes			2 (1.8)
No			108 (98.2)
妊娠前の母体喫煙			
Yes			7 (6.4)
No			103 (93.6)
絨毛膜羊膜炎 (CAM)			
Yes			25 (22.7)
No			85 (77.3)
特発性の前期破水 (iPROM)			
Yes			11 (10.0)
No			99 (90.0)
妊娠高血圧症			
Yes			20 (18.2)
No			90 (81.8)
母体年齢	33.8 (4.7)	34	
<25 歳			1 (0.9)
25 ~ 30 歳			20 (18.2)
30 ~ 35 歳			41 (37.3)
35 ~ 40 歳			34 (30.9)
>40 歳			14 (12.7)
父親の年齢	36.4 (6.7)	35.5	
<25 歳			2 (1.8)
25 ~ 30 歳			11 (10.0)
30 ~ 35 歳			31 (28.2)
35 ~ 40 歳			33 (30.0)
>40 歳			33 (30.0)

	平均値 (SD)	中央値	N (%)
父親の BMI	23.1 (2.9)	22.7	
<18.5 kg/m ²			2 (1.8)
18.5 ~ 25 kg/m ²			82 (74.5)
25 ~ 30 kg/m ²			23 (20.9)
>30 kg/m ²			3 (2.7)
生殖補助医療による妊娠 (ART)			
Yes			25 (22.7)
No			85 (77.3)
母体の妊娠糖尿病 (GDM)			
Yes			4 (3.6)
No			106 (83.5)
前置胎盤			
Yes			10 (9.1)
No			100 (90.9)
経産数			
>0			43 (39.1)
0			67 (60.9)

34.5%が SGA(ここでは出生体重が 10 パーセントイル未満)で、10.9%が large for gestational age (LGA; ここでは出生体重が 90 パーセントイル以上)であった。

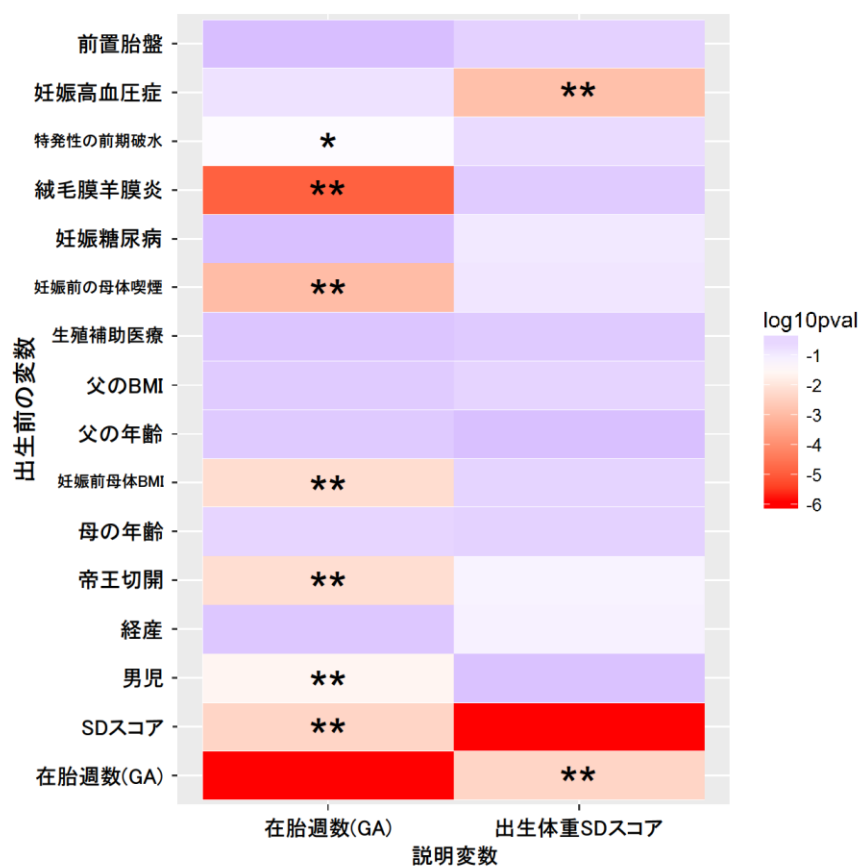
48.2%の児が男児であった。80.9%の児が帝王切開で出生していた。母体の平均年齢は 33.8 歳であった。母体の平均妊娠前 BMI は 21.1 であった。妊娠中の母体喫煙は 2 人のみ(1.8%)であり、そのためこの項目は以降の解析には用いなかった。7 人の母体(6.4%)が妊娠前まで喫煙していた。3.6%(4 人)の母体が妊娠糖尿病、25%が絨毛膜羊膜炎、11%が絨毛膜羊膜炎などの炎症を伴わない特発性の

前期破水、20%が妊娠高血圧症を合併していた。

1. 在胎週数ないし出生体重 SD スコアと関連する臨床的特徴

本研究では出生前の変数と在胎週数および出生体重 SD スコアとの関連について評価した。在胎週数は出生体重 SD スコア、児の性別、出産様式、母体の妊娠前 BMI, 母体の妊娠前喫煙および絨毛膜羊膜炎と関連を認めた (図 12、表 3; $p < 0.05$)。

図 12. 在胎週数ないし出生体重 SD スコアと出生前変数の関係 (n=110, 臍帯血)



*印は各説明変数と各出生前変数の関連度(p 値)を表す。(** = p-value < 0.05 (単変

量回帰分析); * = p-value of 0.060(suggestive))

表 3: 出生前変数と在胎週数の関連—単変量回帰分析の結果

回帰係数 (Est)、95%信頼区間(CI)、*p*-value はそれぞれ、1) 連続出生前変数の場合は、標準偏差の 2 倍の変化に対しての週数変化について評価した結果、2) 二値変数の場合は、2 つのカテゴリーの比較で評価した結果を表す。

	Est	95%CI	<i>p</i> -value
出生体重 SD-score	2.6	(0.83, 4.4)	0.0046
男児	-2.0	(-3.9, -0.21)	0.029
経産	-0.41	(-2.3, 1.5)	0.67
帝王切開	-3.1	(-5.4, -0.86)	0.0075
母体年齢	0.90	(-0.96, 2.8)	0.34
母体妊娠前 BMI	-2.5	(-4.3, -0.70)	0.0071
父の年齢	-0.53	(-2.4, 1.3)	0.58
父の BMI	-0.57	(-2.4, 1.3)	0.54
生殖補助医療 (ART)	0.40	(-1.8, 2.6)	0.72
母体の妊娠前喫煙	-6.1	(-9.7, -2.5)	0.0011
妊娠糖尿病 (GDM)	-0.24	(-5.2, 4.7)	0.93
絨毛膜羊膜炎 (CAM)	-4.7	(-6.7, -2.7)	1.1E-05
特発性の前期破水 (iPROM)	-2.9	(-6.0, 0.12)	0.060
妊娠高血圧症	-1.6	(-4.0, 0.80)	0.19
前置胎盤	-0.0057	(-3.2, 3.2)	0.99

表 4: 出生前変数と出生体重 SD スコアの関連—単変量回帰分析の結果

回帰係数 (Est)、95%信頼区間(CI)、*p*-value はそれぞれ、1) 連続出生前変数の場合は、標準偏差の 2 倍の変化に対しての SD スコア変化について評価した結果、2) 二値変数の場合は、2 つのカテゴリーの比較で評価した結果を表す。

	Est	95%CI	<i>p</i> -value
在胎週数	0.76	(0.24, 1.3)	0.0046
男児	0.058	(-0.48, 0.60)	0.83
経産	-0.46	(-1.0, 0.080)	0.094
帝王切開	-0.59	(-1.3, 0.087)	0.087

	Est	95%CI	p-value
母体年齢	0.23	(-0.31, 0.77)	0.40
母体妊娠前 BMI	0.24	(-0.29, 0.78)	0.37
父の年齢	-0.031	(-0.57, 0.51)	0.91
父の BMI	0.25	(-0.29, 0.79)	0.36
生殖補助医療 (ART)	0.18	(-0.46, 0.82)	0.58
母体の妊娠前喫煙	0.79	(-0.31, 1.9)	0.16
妊娠糖尿病 (GDM)	1.1	(-0.35, 2.5)	0.14
絨毛膜羊膜炎 (CAM)	0.19	(-0.45, 0.83)	0.56
特発性の前期破水 (iPROM)	0.51	(-0.38, 1.4)	0.26
妊娠高血圧症	-1.1	(-1.8, -0.44)	0.0014
前置胎盤	-0.39	(-1.3, 0.55)	0.41

在胎週数は特発性の前期破水とも suggestive な関連を認めた ($p = 0.060$)。出生体重 SD スコアは在胎週数以外に妊娠高血圧症との関連を認めた ($p < 0.05$)。出生体重 SD スコアが高くなるほど、在胎週数は高くなる方向に関連していた (表 3, 表 4)。男児、帝王切開、母体の妊娠前母体 BMI 高値、母体の妊娠前喫煙、絨毛膜羊膜炎は在胎週数が下がる方向に関連していた。特発性の前期破水は在胎週数が下がる方向に suggestive なレベルで関連していた。

GA(在胎週数)、出生体重 SD スコア、児の性別に加え、在胎週数または出生体重 SD スコアに関連する 6 つの出生前の変数を選択し、以降の解析に用いた。6 つの変数は分娩様式、母体の妊娠前 BMI、母体の妊娠前喫煙、絨毛膜羊膜炎、特発性の前期破水、妊娠高血圧症である。これらの変数を用いた多変量解析では

各変数をもたらす効果の方向性としては、単変量解析のときと比較して大きな変化はなかった (表 5、表 6)。しかし、在胎週数と児の性別や母体の妊娠前 BMI との関連は有意ではなくなり、一方で特発性の前期破水や妊娠高血圧症との関連は有意に転じた。

表 5: 出生前変数と在胎週数の関連—多変量回帰分析の結果

回帰係数 (Est)、95%信頼区間(CI)、*p*-value はそれぞれ、1) 連続出生前変数の場合は、標準偏差の 2 倍の変化に対しての週数変化について評価した結果、2) 二値変数の場合は、2 つのカテゴリーの比較で評価した結果を表す。

	Est	95%CI	<i>p</i> -value
出生体重 SD-score	2.6	(1.1, 4.1)	0.0011
男児	-1.1	(-2.6, 0.41)	0.15
帝王切開	-2.5	(-4.4, -0.68)	0.0081
母体妊娠前 BMI	-0.68	(-2.3, 0.97)	0.42
母体の妊娠前喫煙	-4.4	(-7.7, -1.2)	0.0076
絨毛膜羊膜炎 (CAM)	-4.6	(-6.5, -2.7)	6.8E-06
特発性の前期破水 (iPROM)	-4.7	(-7.2, -2.3)	0.00027
妊娠高血圧症	-2.6	(-4.7, -0.49)	0.016

表 6: 出生前変数と出生体重 SD スコアの関連—多変量回帰分析の結果

回帰係数 (Est)、95%信頼区間(CI)、*p*-value はそれぞれ、1) 連続出生前変数の場合は、標準偏差の 2 倍の変化に対しての SD スコア変化について評価した結果、2) 二値変数の場合は、2 つのカテゴリーの比較で評価した結果を表す。

	Est	95%CI	<i>p</i> -value
在胎週数	1.1	(0.44, 1.7)	0.0011
男児	-0.10	(-0.64, 0.43)	0.70
帝王切開	-0.29	(-0.95, 0.38)	0.40
母体妊娠前 BMI	0.50	(-0.062, 1.1)	0.080
母体の妊娠前喫煙	0.78	(-0.37, 1.9)	0.18
絨毛膜羊膜炎 (CAM)	0.40	(-0.33, 1.1)	0.28

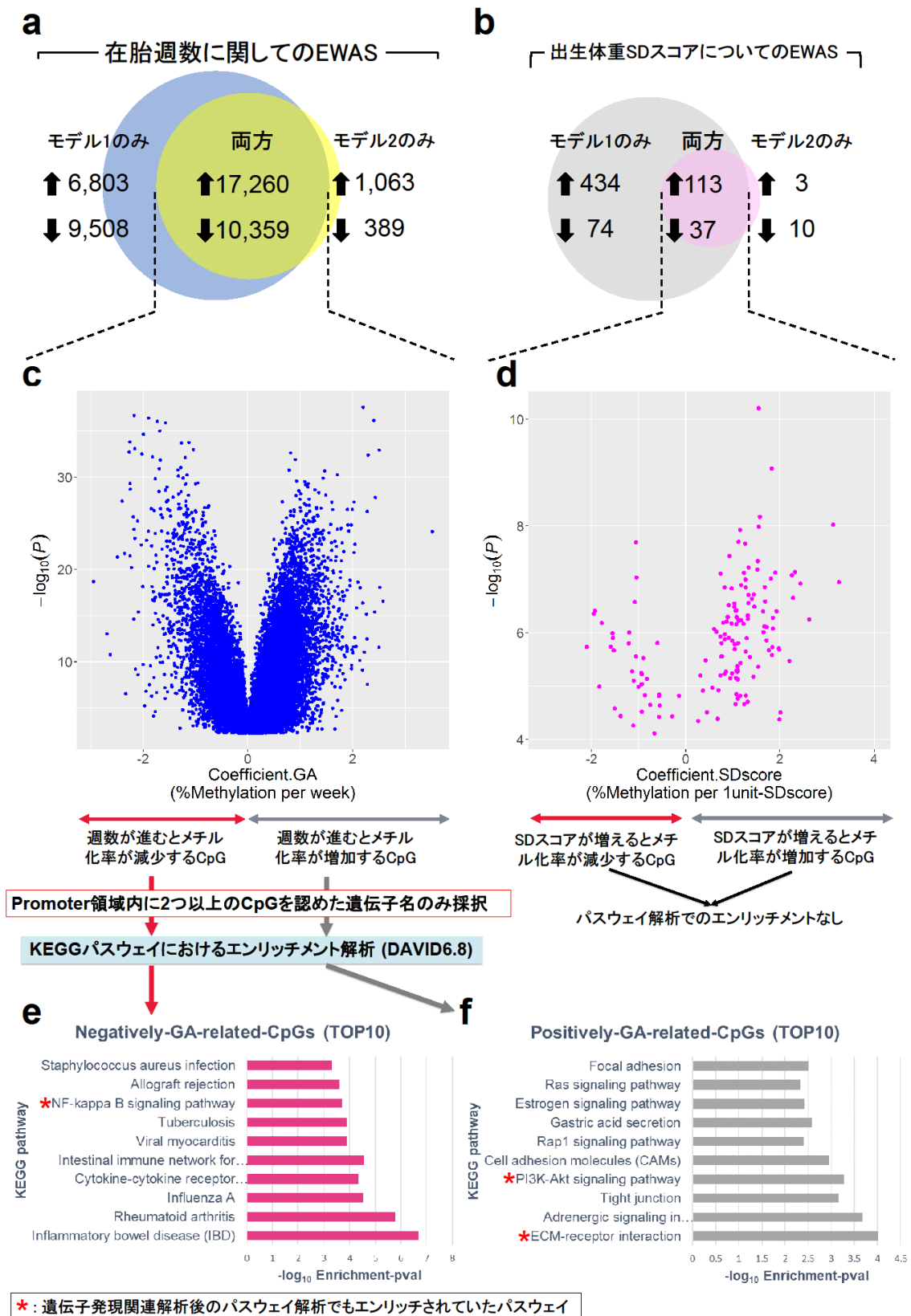
	Est	95%CI	p-value
特発性の前期破水 (iPROM)	0.87	(-0.036, 1.8)	0.060
妊娠高血圧症	-0.87	(-1.6, -0.14)	0.020

2. 在胎週数ないし出生体重 SD スコアについてのエピゲノムワイド関連解析

(EWAS)およびパスウェイ解析

在胎週数ないし出生体重 SD スコアについてのエピゲノムワイド関連解析(EWAS: Epigenome-Wide Association Study)の結果を図 13(次頁)に示す。「モデル 1」(説明変数: 在胎週数および出生体重 SD スコア、目的変数: メチル化率、調整変数: 細胞組成、バッチ、児の性別)による線形回帰分析の結果と「モデル 2」(モデル 1 に調整変数として、絨毛膜羊膜炎の有無、特発性の前期破水の有無、妊娠高血圧症の有無、妊娠前母体喫煙の有無、妊娠前母体 BMI、帝王切開による出生の有無を加えたもの)による結果を比較して回帰係数の正負も含めて共通の CpG サイトを在胎週数関連 CpG サイトおよび SD スコア関連 CpG サイトと決定した(図 13a、b、図 14、図 15)。在胎週数関連 CpG サイトは全部で 27,619 箇所であり、このうちメチル化率増加型(週数が進むとメチル化率が増加するタイプ)は 17,260 箇所、メチル化率減少型は 10,359 箇所であった(図 13a、c)。

図 13. 在胎週数ないし出生体重SDスコアについてのエピゲノムワイド関連解析 (EWAS) (n = 110, 臍帯血)



(図 13 説明) a 青の円は「モデル 1」の線形回帰分析で在胎週数と関連のあった CpG を表し、黄の円は「モデル 2」の線形回帰分析で在胎週数と関連のあった CpG を表す。青と黄の円の交わりの部分が、最終的に決定した「在胎週数関連 CpG」である。**b** 灰色の円は「モデル 1」の線形回帰分析で出生体重 SD スコアと関連のあった CpG を表し、ピンクの円は「モデル 2」の線形回帰分析で SD スコアと関連のあった CpG を表す。灰色とピンクの円の交わりの部分が、最終的に決定した「SD スコア関連 CpG」である。**c** 在胎週数関連 CpG の「モデル 1」における回帰係数を x 軸に、 p -value ($-\log_{10}$ スケール) を y 軸にとった Volcano plot。**d** SD スコア関連 CpG の「モデル 1」における回帰係数を x 軸に、 p -value ($-\log_{10}$ スケール) を y 軸にとった Volcano plot。**e** プロモーター(ここでは TSS1500, TSS200, 5' UTR, 1stExon の領域)上において、2 つ以上メチル化率減少型在胎週数関連 CpG を認めた遺伝子名を用いて行ったパスウェイ解析 (DAVID 6.8) の結果。Enrichment-FDR の上位 10 個のみ提示。**f** プロモーター上において、2 つ以上メチル化率増加型在胎週数関連 CpG を認めた遺伝子名を用いて行ったパスウェイ解析 (DAVID 6.8) の結果。Enrichment-FDR の上位 10 個のみ提示。

*モデル 1) 説明変数: 在胎週数、SD スコア、調整変数: 児の性別、実験バッチ、推定細胞分画

**モデル 2) 説明変数: 在胎週数、SD スコア、調整変数: 児の性別、実験バッチ、推定細胞分画、絨毛膜羊膜炎、特発性の前期破水、妊娠高血圧症、母体の妊娠前喫煙、母体の妊娠前 BMI、帝王切開

***GA: gestational age (在胎週数), EWAS: epigenome-wide association study (エピゲノムワイド関連解析)



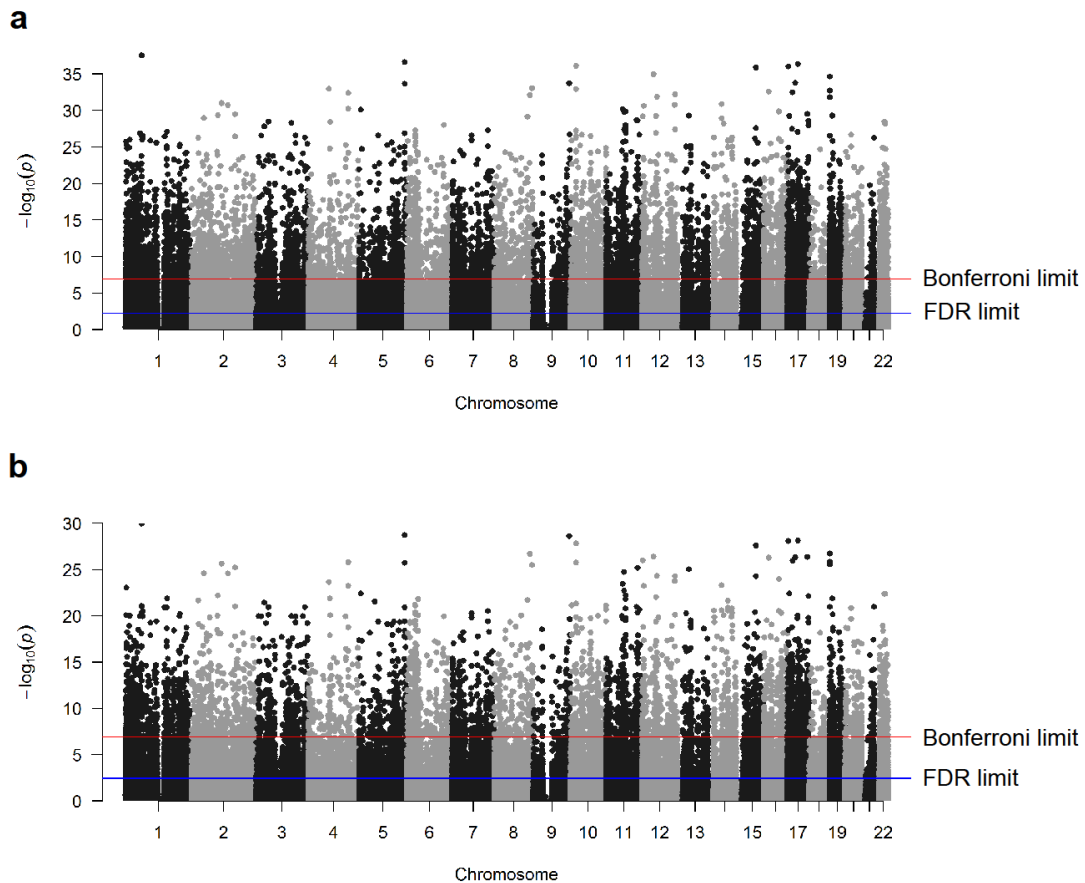
****  (): 説明変数が増加した際にメチル化率が増加する (減少する) CpG。

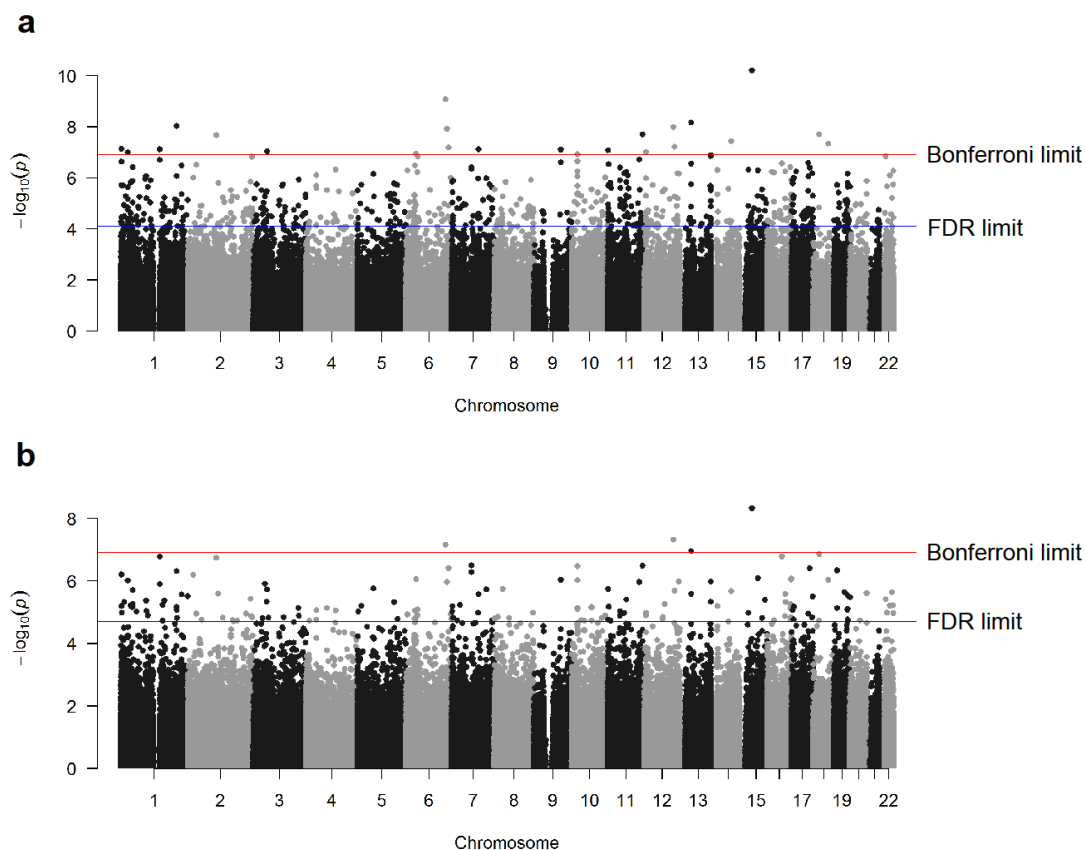
図 14. DNA メチル化と在胎週数の関連度(p -value) についての Manhattan Plot ($n = 110$, 臍帯血)



a 「モデル 1」による線形回帰分析の結果。ゲノムオーダー(染色体番号と染色体上の位置を加味した座標)を x 軸に、 p -value($-\log_{10}$ スケール)を y 軸にとった。水平に走る赤い直線は Bonferroni 補正による閾値(0.05)を、青い直線は FDR 補正による閾値(0.05)を表す。Bonferroni 基準で関連のある CpG は 9,823 箇所、FDR 基準で関連のある CpG は 43,930 箇所であった。**b** 「モデル 2」による線形回帰分析の結果。Bonferroni 基準で関連のある CpG は 5,820 箇所、FDR 基準で関連のある CpG は 29,071 箇所であった。

*「モデル 1, 2」に関しては図 13 参照

図 15. DNA メチル化と出生体重 SD スコアの関連度(p -value) についての Manhattan Plot (n = 110, 臍帯血)



a 「モデル 1」による線形回帰分析の結果。ゲノムオーダー(染色体と染色体上の座標)を x 軸に、 p -value($-\log_{10}$ スケール)を y 軸にとった。水平に走る赤い直線は Bonferroni 補正による閾値(0.05)を、青い直線は FDR 補正による閾値(0.05)を表す。Bonferroni 基準で関連のある CpG は 23 箇所、FDR 基準で関連のある CpG は 658 箇所であった。**b** 「モデル 2」による線形回帰分析の結果。Bonferroni 基準で関連のある CpG は 4 箇所、FDR 基準で関連のある CpG は 163 箇所であった。

*「モデル 1, 2」に関しては図 13 参照

対して、SD スコア関連 CpG サイトは全部で 150 箇所であり、メチル化率増加型(SD スコアが増えるとメチル化率が増加するタイプ)は 113 箇所、メチル化率減少型は 37 箇所であった(図 13b, d)。

パスウェイ解析の結果は、SD スコア関連 CpG サイトに関してはメチル化率増加型・減少型どちらのカテゴリーでも enrichment-FDR \leq 0.1 を満たす KEGG パスウェイはなかった。在胎週数関連 CpG サイトに関しては、メチル化率増加型のものでは ECM-receptor interaction, PI3K-Akt signaling pathway の他、Adrenergic signaling in cardiomyocytes, Estrogen signaling pathway などが検出され、FDR $>$ 0.1 であるものの suggestive なレベルのものとして、Oxytocin signaling pathway, Signaling pathways regulating pluripotency of stem cells が検出された(図 13e, 表 7)。メチル化率減少型のものには Cytokine-cytokine receptor interaction や NF-kappa B signaling pathway などの pathway が検出された。

表 7: 在胎週数関連 CpG における、KEGG パスウェイでのエンリッチメント解析結果

*ER-*p*-value: DAVID エンリッチメント解析で算出された enrichment-*p*-value

**ER-FDR: Benjamini-Hochberg 法による多重検定補正後の enrichment-*p*-value

GA: gestational age *: ER-FDR $>$ 0.1 であるが、suggestive な関連性。

[Positively-GA-related-CpGs (週数が進むとメチル化率が増加するタイプ)]

KEGG pathway	Count	ER- <i>p</i> -Value*	ER-FDR**
ECM-receptor interaction	16	9.8E-05	0.025
Adrenergic signaling in cardiomyocytes	21	2.1E-04	0.026

KEGG pathway	Count	ER- <i>p</i> -Value*	ER-FDR**
Tight junction	19	7.0E-04	0.044
PI3K-Akt signaling pathway	36	5.3E-04	0.045
Cell adhesion molecules (CAMs)	19	1.1E-03	0.054
Rap1 signaling pathway	23	3.9E-03	0.095
Gastric acid secretion	12	2.6E-03	0.10
Estrogen signaling pathway	14	3.8E-03	0.10
Ras signaling pathway	24	4.6E-03	0.10
Focal adhesion	23	3.1E-03	0.11****
Oxytocin signaling pathway	19	3.6E-03	0.11****
Signaling pathways regulating pluripotency of stem cells	17	5.7E-03	0.12****

[Negatively-GA-related-CpGs (週数が進むとメチル化率が増加するタイプ)]

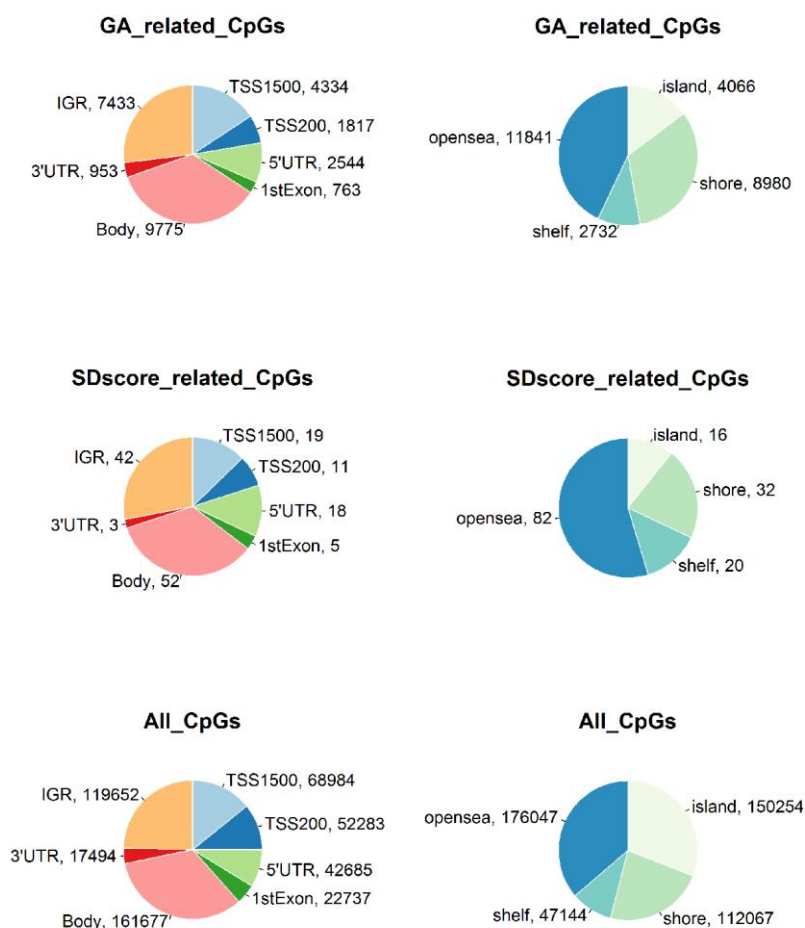
KEGG pathway	Count	ER- <i>p</i> -Value*	ER-FDR**
Inflammatory bowel disease (IBD)	14	2.1E-07	5.0E-05
Rheumatoid arthritis	15	1.7E-06	2.0E-04
Influenza A	19	3.0E-05	1.8E-03
Cytokine-cytokine receptor interaction	22	4.4E-05	2.1E-03
Intestinal immune network for IgA production	10	2.7E-05	2.2E-03
Viral myocarditis	10	1.3E-04	4.5E-03
Tuberculosis	18	1.3E-04	5.1E-03
NF-kappa B signaling pathway	12	1.9E-04	5.7E-03
Allograft rejection	8	2.4E-04	6.3E-03
Staphylococcus aureus infection	9	4.9E-04	0.012
HTLV-I infection	21	5.5E-04	0.012
Herpes simplex infection	17	5.9E-04	0.012
Graft-versus-host disease	7	8.3E-04	0.015
Primary immunodeficiency	7	9.8E-04	0.017
Chemokine signaling pathway	16	2.0E-03	0.029
Measles	13	2.2E-03	0.030
Autoimmune thyroid disease	8	2.0E-03	0.031
Type I diabetes mellitus	7	3.0E-03	0.036
T cell receptor signaling pathway	11	2.9E-03	0.038

KEGG pathway	Count	ER- <i>p</i> -Value*	ER-FDR**
Leishmaniasis	9	3.0E-03	0.038
Asthma	6	3.4E-03	0.038
Jak-STAT signaling pathway	13	4.4E-03	0.047
Malaria	7	6.6E-03	0.067
Toxoplasmosis	11	7.7E-03	0.075
Natural killer cell mediated cytotoxicity	11	9.7E-03	0.089
Cell adhesion molecules (CAMs)	12	0.010	0.091
Amoebiasis	10	0.011	0.095

在胎週数関連 CpG サイトと SD スコア関連 CpG サイトの所在についても検討したが、TSS(Transcription Starting Site: 転写開始点)を中心としたゲノム領域分類(TSS1500: 転写開始点から-1500 ~ -200 塩基の領域、TSS200: 転写開始点から-200 塩基までの領域、5'UTR: 5'側の非翻訳領域、1stExon: 第1エクソン、Body: gene body、3'UTR: 3'側の非翻訳領域、IGR: 遺伝子間領域)では、アレイ内 CpG の所在配分と大きく変わらなかった(図 16; 次頁)。CpG アイランドを中心とした分類(CpG island: 200 塩基以上の長さを有し、GC 部分が全体の 50%以上を占める領域、CpG shore: CpG アイランドの一端から 2,000 塩基の間の領域、CpG shelf: CpG アイランドの一端から 2000 ~ 4000 塩基の領域、Open Sea: 左のどのカテゴリーにも属さず、GC 部分の少ない領域)では、在胎週数関連 CpG サイトでは CpG shore が 8,980 箇所(32.5%; アレイ内全 CpG は 22.9%)と最も多く、SD スコア関

連 CpG サイトでは Open Sea が 82 箇所(54.7%; アレイ内全 CpG は 36.0%)と最も多かった。

図 16. 在胎週数関連 CpG ないし出生体重 SD スコア関連 CpG の分布



左右の区分：左のパネルがゲノム上 TSS (転写開始点) を中心としたゲノム領域分類の分布で、右のパネルが CpG アイランドを中心とした分類の分布。上中下の区分：上のパネルが在胎週数関連 CpG についてのもので、中のパネルが SD スコア関連 CpG についてのもの、下のパネルが DNA メチル化アレイ HumanMethylation450 BeadChip に含まれる CpG 全体についてのもの。

*GA: gestational age(在胎週数), SD-score: birthweight SD-score(出生体重 SD スコア)

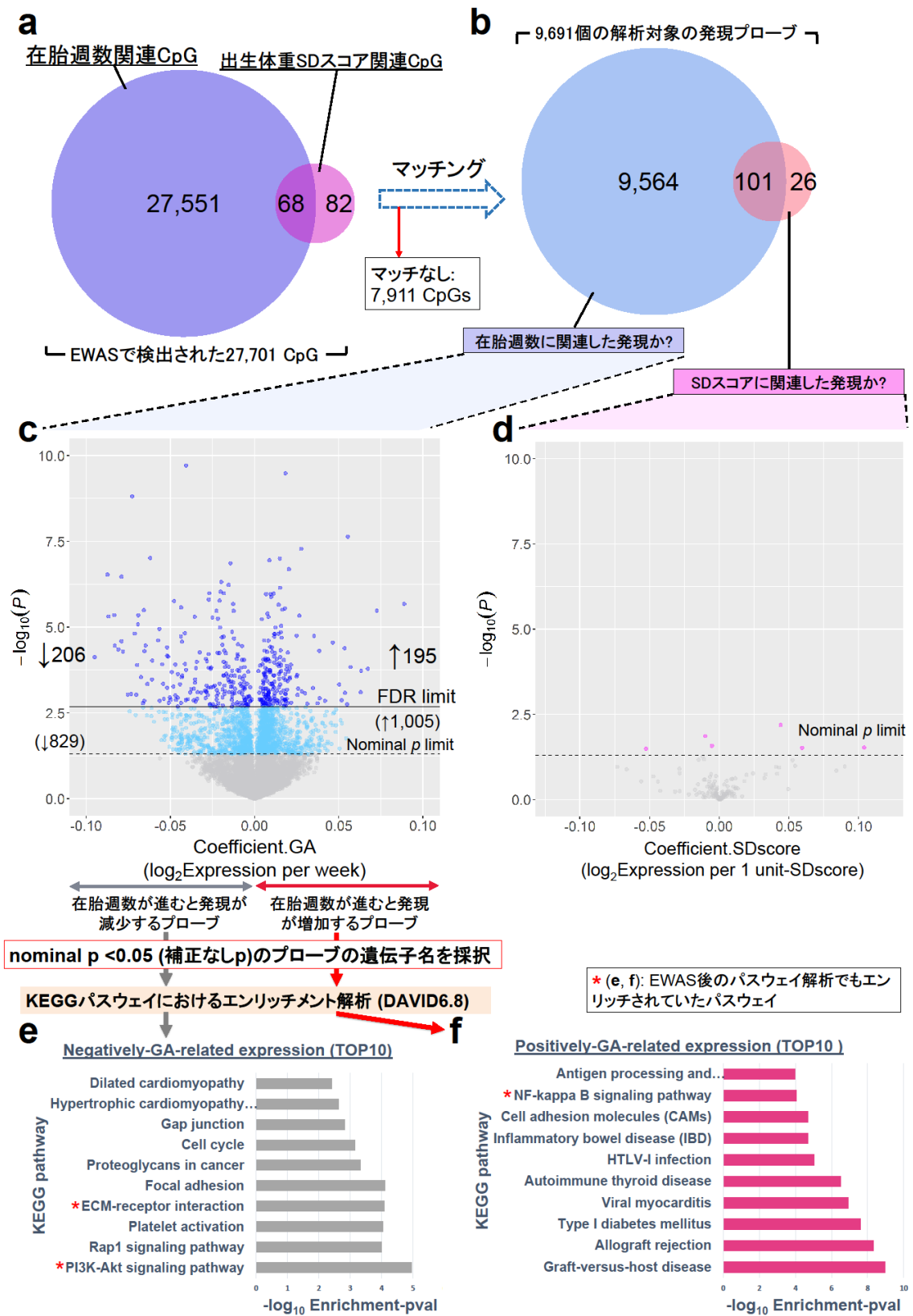
3. 在胎週数ないし出生体重 SD スコアについての遺伝子発現関連解析および

パスウェイ解析

在胎週数ないし出生体重 SD スコアについての EWAS で決定した関連 CpG サイト 27,701 箇所のうち 19,790 箇所に対応する 9,691 種類の遺伝子発現プローブについて「モデル 1」(説明変数: 在胎週数および出生体重 SD スコア、目的変数: $\log_2\text{Expression}$ (全サンプル全プローブ中発現シグナル強度が最小のものを 1 に設定し、シグナル強度に対し底 2 の対数変換を行ったもの)、調整変数: 細胞組成、バッチ、児の性別)による線形回帰分析を用いて行った、遺伝子発現関連解析の結果を図 17c, d に示す。SD スコア関連遺伝子発現プローブは補正なしの p 値が 0.05 未満のものは 6 個認めたが、多重検定補正後 $\text{FDR}<0.05$ を満たすプローブは検出しなかった(図 17d)。在胎週数関連のものに関しては $\text{FDR}<0.05$ を満たすプローブを計 401 個(週数が進むと発現が増加するプローブが 195 個、対して発現減少型が 206 個)、[補正なし p 値] <0.05 を満たすプローブを計 1,834 個(発現増加型が 1,005 個で、発現減少型が 829 個)認めた(図 17c)。後者はパスウェイ解析にのみ用いて、前者を「在胎週数関連遺伝子発現プローブ」と決定しパスウェイ解析を含む以降の解析に用いた。

(本文は図 17 を間に挟んで続く。)

図 17. 在胎週数ないし出生体重 SD スコアについての遺伝子発現関連解析 (n = 55, 臍帯血)



(図17 説明) **a** 青い円は臍帯血のEWASで在胎週数と関連のあったCpGを表し、ピンクの円は出生体重SDスコアと関連のあったCpGを表す。 **b** 臍帯血のEWASで検出された全27,701箇所のCpGサイトのうち、19,640箇所を9,691個の遺伝子発現プローブにマッチした。水色の円は在胎週数関連CpGに対応した遺伝子発現プローブを表し、ピンクの円はSDスコア関連CpGに対応した発現プローブを表す。 **c** 在胎週数についての遺伝子発現解析結果のVolcano plot。x軸は回帰係数、y軸はp-value (-log10スケール)を表す。解析には、EWASでも用いた「モデル1」による線形回帰分析を応用した。濃い青のドットはFDR基準で関連を認めたプローブを表す。水色のドットはFDR基準を満たさないものの、補正なしのp値が0.05未満のプローブを表す。灰色のドットは補正なしのp値でも関連を認めなかったプローブを表す。 **d** 出生体重SDスコアについての遺伝子発現解析結果のVolcano plot。x軸は回帰係数、y軸はp-value (-log10スケール)を表す。解析には、EWASでも用いた「モデル1」による線形回帰分析を応用した。ピンクのドットはFDR基準を満たさないものの、補正なしのp値が0.05未満のプローブを表す。灰色のドットは補正なしのp値でも関連を認めなかったプローブを表す。 **e** [補正なしp値]<0.05を満たす発現減少型の遺伝子発現プローブを用いて行ったパスウェイ解析の結果。Enrichment-FDRの上位10個のみ提示。棒グラフのx軸はenrichment-p-value (-log10スケール)を表す。 **f** [補正なしp値]<0.05を満たす発現増加型の遺伝子発現プローブを用いて行ったパスウェイ解析の結果。Enrichment-FDRの上位10個のみ提示。棒グラフのx軸はenrichment-p-value (-log10スケール)を表す。

*モデル1) 説明変数: 在胎週数、出生体重SDスコア、目的変数: 遺伝子発現レベル(全サンプル全プローブの中で信号強度が最小のものを1として表した発現レベルをlog2変換したもの)、調整変数: 児の性別、実験バッチ、細胞組成

**GA: gestational age(在胎週数), SD-score: birthweight SD-score(出生体重SDスコア), EWAS: epigenome-wide association study (エピゲノムワイド関連解析)

在胎週数関連遺伝子発現プローブ(FDR<0.05 での採択)を用いたパスウェイ解析では発現増加型では enrichment-FDR \leq 0.1 をみたす KEGG パスウェイは検出されず、発現増加型プローブは Platelet activation のみの検出であった(表 8)。

次に、[補正なし p 値]<0.05 を満たすプローブを用いて再度、パスウェイ解析を行ったところ、発現増加型では NF-kappa B signaling pathway、Cytokine-cytokine interaction など、発現減少型では PI3K-Akt signaling pathway、ECM-receptor interaction などのパスウェイが検出された(図 17e, f, 表 9)。

表 8: FDR 基準で採択した在胎週数関連発現プローブ**における、KEGG パスウェイでのエンリッチメント解析結果**

ここでは、enrichment p -value \leq 0.1 の KEGG パスウェイを掲示している。

*ER- p -value: DAVID エンリッチメント解析で算出された enrichment- p -value

**ER-FDR: Benjamini-Hochberg 法による多重検定補正後の enrichment- p -value

***GA: gestational age

[Negatively-GA-related expression probes(週数が進むと発現が減少するプローブ)]

KEGG pathway	Count	ER- p -Value*	ER-FDR**
Platelet activation	9	4.5E-05	6.3E-03
Rap1 signaling pathway	8	5.1E-03	0.30
Pathways in cancer	11	5.8E-03	0.24
Cell cycle	6	8.2E-03	0.25
Focal adhesion	7	0.018	0.39
MicroRNAs in cancer	8	0.025	0.44
Asthma	3	0.037	0.53
Proteoglycans in cancer	6	0.051	0.60
TGF-beta signaling pathway	4	0.053	0.57
Hematopoietic cell lineage	4	0.055	0.54
ECM-receptor interaction	4	0.058	0.53

KEGG pathway	Count	ER-p-Value*	ER-FDR**
PI3K-Akt signaling pathway	8	0.059	0.51
Prostate cancer	4	0.059	0.48
Gap junction	4	0.059	0.48
Cytokine-cytokine receptor interaction	6	0.083	0.58
Transcriptional misregulation in cancer	5	0.090	0.59

[Positively-GA-related expression probes(週数が進むと発現が増加するプローブ)]

KEGG pathway	Count	ER-p-Value*	ER-FDR**
Epstein-Barr virus infection	5	0.048	1
Phagosome	4	0.098	1

表 9: [補正なし p 値] <0.05 の基準で採択した在胎週数関連発現プローブにおける、KEGG パスウェイでのエンリッチメント解析結果

*ER-p-value: DAVID エンリッチメント解析で算出された enrichment-p-value

**ER-FDR: Benjamini-Hochberg 法による多重検定補正後の enrichment-p-value

***GA: gestational age

[Negatively-GA-related expression probes(週数が進むと発現が減少するプローブ)]

KEGG pathway	Count	ER-p-Value*	ER-FDR**
PI3K-Akt signaling pathway	31	1.1E-05	2.6E-03
Rap1 signaling pathway	21	1.0E-04	4.8E-03
Platelet activation	16	9.1E-05	5.4E-03
ECM-receptor interaction	13	8.3E-05	6.5E-03
Focal adhesion	21	7.9E-05	9.2E-03
Proteoglycans in cancer	19	4.6E-04	0.018
Cell cycle	14	6.9E-04	0.023
Gap junction	11	1.5E-03	0.044
Hypertrophic cardiomyopathy (HCM)	10	2.3E-03	0.060
Dilated cardiomyopathy	10	3.9E-03	0.088

[Positively-GA-related expression probes(週数が進むと発現が増加するプローブ)]

KEGG pathway	Count	ER-p-Value*	ER-FDR**
Graft-versus-host disease	13	1.1E-09	2.5E-07
Allograft rejection	13	4.8E-09	5.8E-07
Type I diabetes mellitus	13	2.4E-08	1.9E-06

KEGG pathway	Count	ER-p-Value*	ER-FDR**
Viral myocarditis	14	1.2E-07	7.2E-06
Autoimmune thyroid disease	13	3.2E-07	1.6E-05
HTLV-I infection	26	9.3E-06	3.8E-04
Inflammatory bowel disease (IBD)	12	2.1E-05	6.4E-04
Cell adhesion molecules (CAMs)	18	2.0E-05	6.9E-04
NF-kappa B signaling pathway	13	8.6E-05	2.3E-03
Antigen processing and presentation	12	1.1E-04	2.6E-03
Herpes simplex infection	19	1.6E-04	3.5E-03
Leishmaniasis	11	2.8E-04	5.6E-03
Hematopoietic cell lineage	12	3.0E-04	5.6E-03
Rheumatoid arthritis	12	4.1E-04	7.0E-03
Asthma	7	7.3E-04	0.012
Intestinal immune network for IgA production	8	1.7E-03	0.025
Tuberculosis	16	2.5E-03	0.035
Staphylococcus aureus infection	8	3.8E-03	0.049
Epstein-Barr virus infection	16	5.0E-03	0.061
Cytokine-cytokine receptor interaction	18	5.6E-03	0.065
Signaling pathways regulating pluripotency of stem cells	13	6.1E-03	0.068
Insulin resistance	11	7.0E-03	0.074
Osteoclast differentiation	12	9.9E-03	0.099
Phagosome	13	0.012	0.11****
Toxoplasmosis	11	0.013	0.12****
B cell receptor signaling pathway	8	0.014	0.12****

****: FDR >0.1, suggestive

本項の最後に、確認・参考のため、遺伝子発現解析に用いた臍帯血 55 サンプル

の臨床データの分布を以下、表 10 に示す。平均の在胎週数は 33.0 週で平均の

SD スコアは-0.9 であった。38.2%が SGA(ここでは出生体重が 10 パーセントイ

ル未満)で、5.0%が large for gestational age (LGA; ここでは出生体重が 90 パーセントイル以上)であった。56.4%の児が男児であった。78.2%の児が帝王切開で出生していた。母体の平均年齢は 33.4 歳であった。母体の平均妊娠前 BMI は 21.1 であった。妊娠中の母体喫煙は 2 人(3.6%)であった。5 人の母体(9.1%)が妊娠前まで喫煙していた。1.8%(1 人)の母体が妊娠糖尿病、27.3%が絨毛膜羊膜炎、10.9%が絨毛膜羊膜炎などの炎症を伴わない特発性の前期破水、20.0%が妊娠高血圧症を合併していた。

表 10. 遺伝子発現解析に用いた臍帯血 55 サンプルの臨床データの分布

	平均値 (SD)	中央値	N (%)
在胎週数	33.0 (5.2)	33.3	
<28 週			10 (18.2)
28-32 週			14 (25.5)
32-37 週			14 (25.5)
>37 週			17 (30.9)
出生体重 SD スコア	-0.9 (1.6)	-0.9	
<-2.5			11 (20.0)
-2.5 ~ -1.28			10 (18.2)
-1.28 ~ 1.28			29 (52.7)
1.28 ~ 2.5			5 (9.1)
性別			
男児			31 (56.4)
女児			24 (43.6)
分娩様式			
帝王切開			43 (78.2)
経膣			12 (21.8)

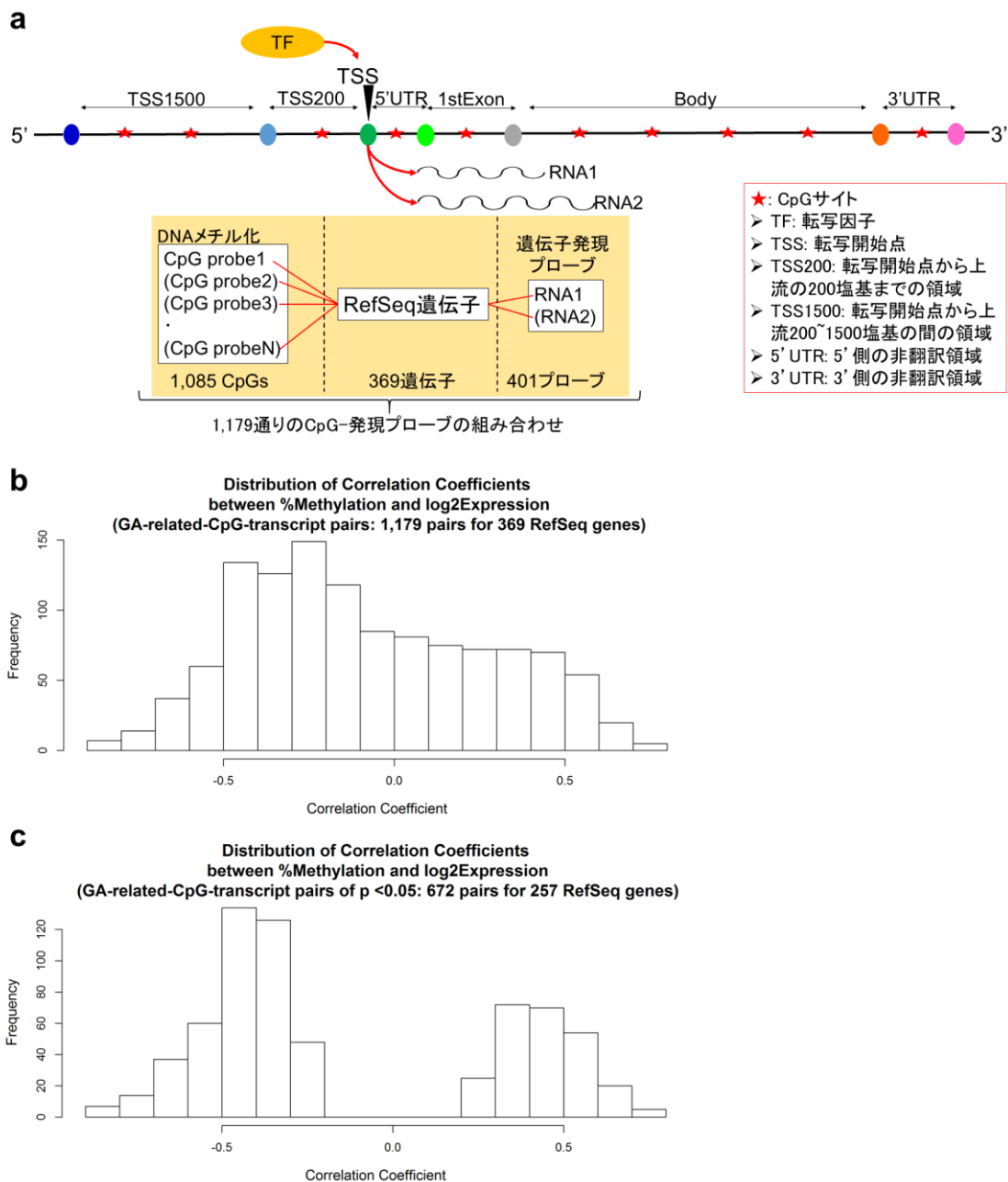
	平均値 (SD)	中央値	N (%)
母体妊娠前 BMI	21.1 (3.9)		20.0
<18.5 kg/m ²			13 (23.6)
18.5 ~ 25 kg/m ²			35 (63.6)
25 ~ 30 kg/m ²			4 (7.3)
>30 kg/m ²			3 (5.5)
妊娠中の母体喫煙			
Yes			2 (3.6)
No			53 (96.4)
妊娠前の母体喫煙			
Yes			5 (9.1)
No			50 (90.9)
絨毛膜羊膜炎 (CAM)			
Yes			15 (27.3)
No			40 (72.7)
特発性の前期破水 (iPROM)			
Yes			6 (10.9)
No			49 (89.1)
妊娠高血圧症			
Yes			11 (20.0)
No			44 (80.0)
母体年齢	33.4 (5.1)		33
<25 歳			1 (1.8)
25 ~ 30 歳			12 (21.8)
30 ~ 35 歳			18 (32.7)
35 ~ 40 歳			17 (30.9)
>40 歳			7 (12.7)
父親の年齢	36.4 (7.7)		35
<25 歳			1 (1.8)
25 ~ 30 歳			7 (12.7)
30 ~ 35 歳			15 (27.3)
35 ~ 40 歳			14 (25.5)
>40 歳			18 (32.7)

	平均値 (SD)	中央値	N (%)
父親の BMI	23.7 (3.4)		23.3
<18.5 kg/m ²			2 (3.6)
18.5 ~ 25 kg/m ²			38 (69.1)
25 ~ 30 kg/m ²			12 (21.8)
>30 kg/m ²			3 (5.5)
生殖補助医療による妊娠 (ART)			
Yes			13 (23.6)
No			42 (76.4)
母体の妊娠糖尿病 (GDM)			
Yes			1 (1.8)
No			54 (98.2)
前置胎盤			
Yes			3 (5.5)
No			52 (94.5)
経産数			
>0			22 (40.0)
0			33 (60.0)

4. DNA メチル化と遺伝子発現の関連の確認

在胎週数関連遺伝子発現プローブ 401 個と対応する在胎週数関連 CpG サイト 1,085 箇所の組み合わせ 1,179 組(RefSeq 遺伝子として 369 個)に対して、メチル化－遺伝子発現間の関連について単変量回帰モデルを用いて解析した結果(相関係数の分布)を図 18(次頁)に示す。

図 18. 在胎週数関連遺伝子発現プローブと対応する在胎週数関連 CpG の DNA メチル化-遺伝子発現間の関連についての検討 (n = 55, 臍帯血)



a. (図 11 の再掲) 1,085 箇所 of 在胎週数関連 CpG サイトと対応する 401 個の在胎週数関連遺伝子発現プローブから 369 個の RefSeq 遺伝子を介して、1,179 通りの直接比較用の CpG-発現プローブの組み合わせを作成。 **b.** 在胎週数関連の CpG と対応する発現プローブの組み合わせ全ての DNA メチル化-遺伝子発現 (log2 スケール)間相関係数の分布 **c.** p<0.05 の組み合わせを除いた場合の、在胎

週数関連の CpG と対応する発現プローブの組み合わせ全ての DNA メチル化-発現(log₂ スケール)間相関係数の分布

*GA: gestational age

1,179 組の DNA メチル化-遺伝子発現間の相関係数の中央値は-0.14 で、相関係数が負の組み合わせが 730 組で、正の組み合わせが 449 組であった(図 18b)。全組み合わせ 1179 組のうち、 $p < 0.05$ の組み合わせのみ採択したところ、672 組の組み合わせ(257 個の RefSeq 遺伝子に相当)がこの基準を満たした。この 672 組のうち、相関係数が負の組み合わせ(メチル化率が増加すると遺伝子発現が減少する組み合わせ)が 426 組で、正の組み合わせが 246 組であった(図 18c)。在胎週数関連 CpG サイト 27,619 箇所のうち、対応する遺伝子に在胎週数との関連を認め、DNA メチル化と遺伝子発現の直接の相関を認めたものは 628 箇所となる。

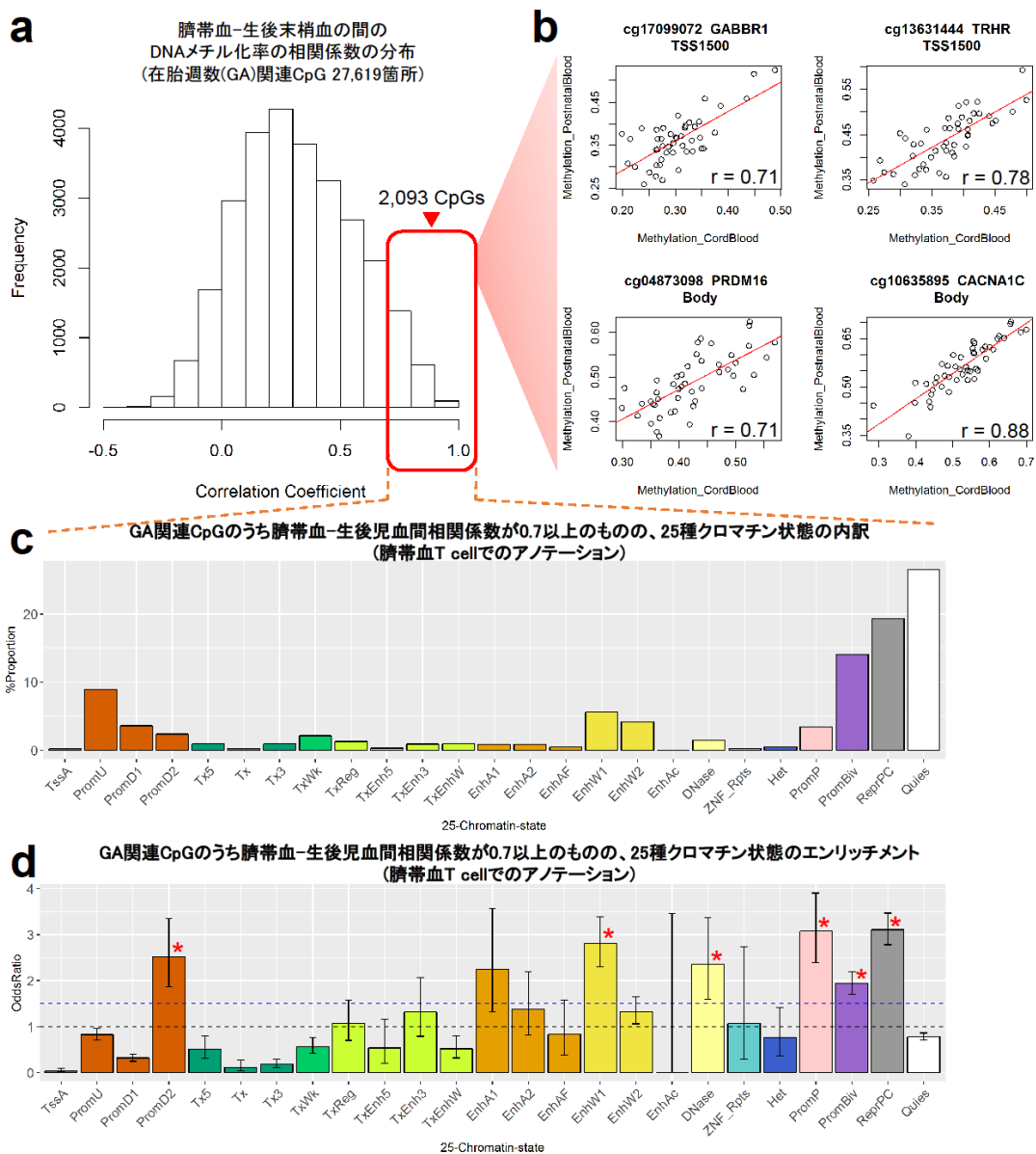
5. エピジェネティックメモリーの候補

在胎週数関連 CpG サイト 27,619 箇所と出生体重 SD スコア関連 CpG サイト 150 箇所の臍帯血-生後末梢血の間の DNA メチル化率の相関係数の分布を図 19a(次頁)、図 20a に示す。在胎週数関連 CpG サイトで相関係数が 0.7 以上のものは 2,093 箇所、SD スコア関連 CpG サイトで相関係数が 0.7 以上のものは 7 箇所であった。SD スコア関連 CpG サイト 7 箇所全ておよび在胎週数関連 CpG サイトの一部(4 箇所)の、臍帯血メチル化率と生後末梢血メチル化率の関係を示す散布図をそれぞれ図 20b、図 19b に示す。

次に在胎週数関連 CpG サイトの中のエピジェネティックメモリーの候補サイトについて、ChromHMM における 12 種類のヒストンマークで特徴づけられる 25 種類のクロマチン状態のエンリッチメントを検討した解析結果について述べる。T cell でのアノテーション(以下 T-cell-Anno)、B cell でのアノテーション(以下 B-cell-Anno)ともに多数認めるクロマチン状態は第 1 位から第 4 位まで同じで、第 1 位 Quiescent/Low (ヒストンマークの関与が少ない状態; T-cell-Anno: 26.5%, B-cell-Anno: 24.2%)、第 2 位 Repressed Polycomb (抑制性のポリコム; T-cell-Anno: 19.3%, B-cell-Anno: 20.8%)、

(本文は図 19、20、21、表 11 を間に挟んで続く。)

図 19. 在胎週数関連のエピジェネティックメモリーの候補 CpG (n = 47, 臍帯血および生後末梢血)

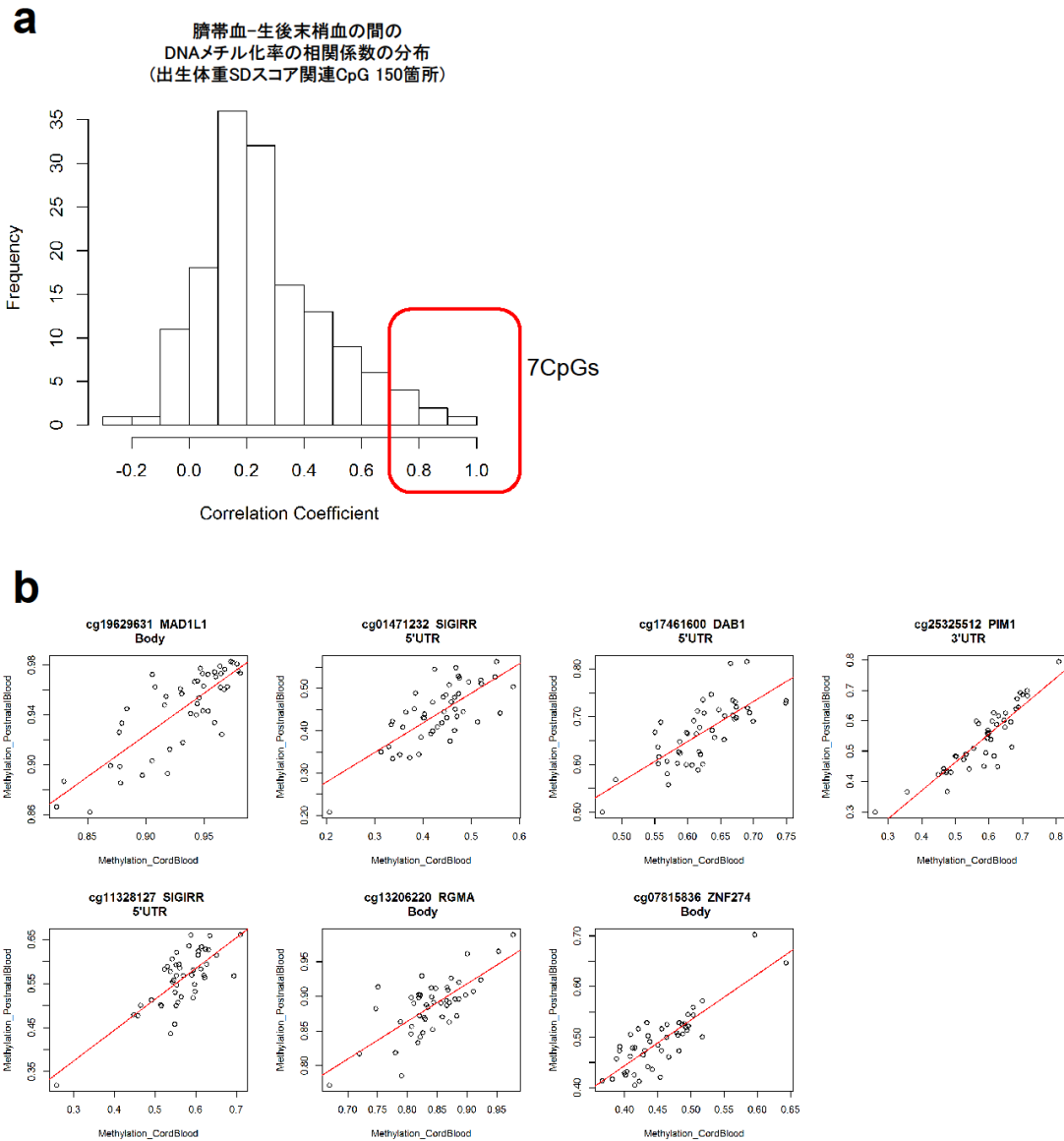


a 在胎週数(GA)関連 CpG27,619 箇所における臍帯血-生後末梢血の間の DNA メチル化率の相関係数の分布を表すヒストグラム。赤線の囲みは相関係数が 0.7 以上、すなわち DNA メチル化変化が生後も持続する CpG の候補である。 **b** 候補 CpG のうち 4 箇所について、x 軸に臍帯血での DNA メチル化率、y 軸に生後末梢血でのメチル化率を表した散布図を例示。 **c** GA(在胎週数)関連 CpG のうち臍帯血-生後児血間相関係数が 0.7 以上のものの、25 種クロマチン状態の内訳 (臍帯血 T cell でのアノテーション)。 **d** GA 関連 CpG のうち臍帯血-生後児血間相関

係数が0.7以上のものの、25種クロマチン状態のエンリッチメント (臍帯血 T cell でのアノテーション)。エラーバーは 95%CI 信頼区間(CI)を表す。*はエンリッチされているクロマチン状態、すなわち Bonferroni 基準 (p-value <0.05/25)で有意な関連がありオッズ比が 1(黒の破線)以上の状態を表す。結果として、エンリッチされていたクロマチン状態の 95%CI の下限は全て 1.5(青の破線)を超えていた。
*GA: gestational age(在胎週数), EWAS: epigenome-wide association study(エピゲノムワイド関連解析)

以下の略語は ChromHMM で定められているものである。: TssA: Active TSS, PromU: Promoter upstream TSS, PromD1: Promoter downstream with DNase, PromD2: Promoter downstream TSS, Tx5: Transcription 5', Tx: Transcription, Tx3: Transcription 3', TxWk: Weak Transcription, TxReg: Transcription Regulatory, TxEnh5: Transcription 5' Enhancer, TxEnh3: Transcription 3' Enhancer, TxEnhW: Transcription Weak Enhancer, EnhA1: Active Enhancer 1, EnhA2: Active Enhancer 2, EnhAF: Active Enhancer Flank, EnhW1: Weak Enhancer 1, EnhW2: Weak Enhancer 2, EnhAc: Enhancer Acetylation Only, DNase: DNase only, ZNF_Rpts: ZNF genes & repeats, Het: Heterochromatin, PromP: Poised Promoter, PromBiv: Bivalent Promoter, ReprPC: Repressed Polycomb, Qies: Quiescent/Low

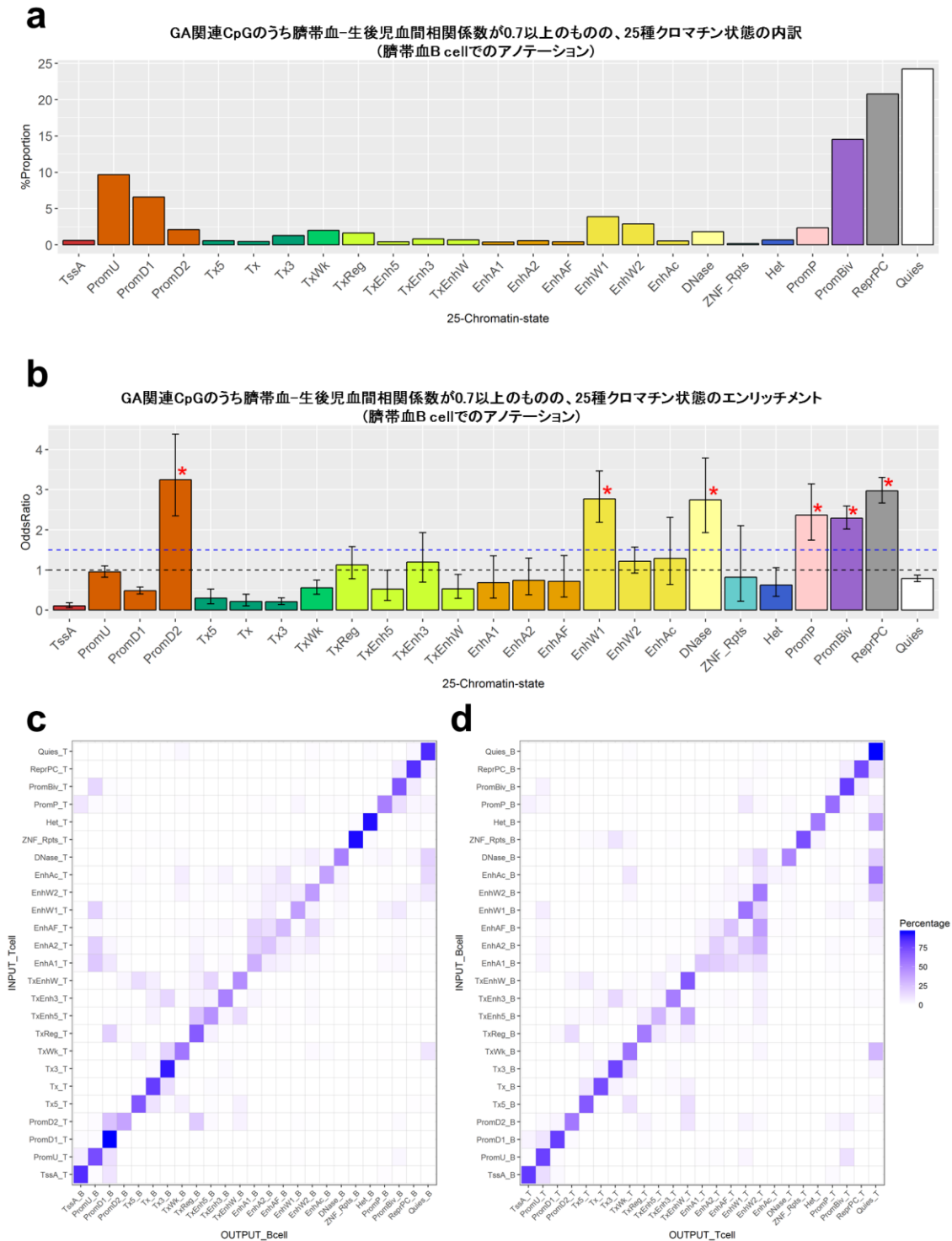
図 20. 出生体重 SD スコア関連のエピジェネティックメモリーの候補 CpG(n = 47, 臍帯血および生後末梢血)



a 出生体重 SD スコア関連 CpG150 箇所における臍帯血-生後末梢血の間の DNA メチル化率の相関係数の分布を表すヒストグラム。赤線の囲みは相関係数が 0.7 以上、すなわち DNA メチル化変化が生後も持続する CpG の候補である。**b** 候補 CpG 全 7 箇所について、x 軸に臍帯血での DNA メチル化率、y 軸に生後末梢血でのメチル化率を表した散布図。

*GA: gestational age(在胎週数), EWAS: epigenome-wide association study(エピゲノムワイド関連解析)

図 21. 在胎週数関連のエピジェネティックメモリーの候補 CpG についての補足 (n = 47, 臍帯血および生後末梢血)



a GA(在胎週数)関連 CpG のうち臍帯血-生後児血間相関係数が 0.7 以上のものの、25 種クロマチン状態の内訳 (臍帯血 T cell でのアノテーション)。**b** GA 関連 CpG

のうち臍帯血-生後児血間相関係数が 0.7 以上のものの、25 種クロマチン状態のエンリッチメント (臍帯血 T cell でのアノテーション)。エラーバーは 95%CI 信頼区間(CI)を表す。*はエンリッチされているクロマチン状態、すなわち Bonferroni 基準 (p-value <0.05/25)で有意な関連がありオッズ比が 1(黒の破線)以上の状態を表す。結果として、エンリッチされていたクロマチン状態の 95%CI の下限は全て 1.5(青の破線)を超えていた。 **c** ChromHMM における、臍帯血 T cell での 25 種クロマチン状態のアノテーションと B cell でのアノテーションの関係性。y 軸が INPUT で、x 軸が OUTPUT。DNA メチル化アレイ上の一つの CpG に、T cell における一つのアノテーションが割り当てられた場合(INPUT)、どの B cell でのアノテーションがどの程度の率で該当するか(OUTPUT)を表したヒートマップ。 **d** 前述 **c** における検討の順序を逆にしたもの。DNA メチル化アレイ上の一つの CpG に、B cell における一つのアノテーションが割り当てられた場合(INPUT)、どの T cell でのアノテーションがどの程度の率で該当するか(OUTPUT)を表したヒートマップ。

*GA: gestational age(在胎週数), EWAS: epigenome-wide association study(エピゲノムワイド関連解析)

以下の略語は ChromHMM で定められているものである。: TssA: Active TSS, PromU: Promoter upstream TSS, PromD1: Promoter downstream with DNase, PromD2: Promoter downstream TSS, Tx5: Transcription 5', Tx: Transcription, Tx3: Transcription 3', TxWk: Weak Transcription, TxReg: Transcription Regulatory, TxEnh5: Transcription 5' Enhancer, TxEnh3: Transcription 3' Enhancer, TxEnhW: Transcription Weak Enhancer, EnhA1: Active Enhancer 1, EnhA2: Active Enhancer 2, EnhAF: Active Enhancer Flank, EnhW1: Weak Enhancer 1, EnhW2: Weak Enhancer 2, EnhAc: Enhancer Acetylation Only, DNase: DNase only, ZNF_Rpts: ZNF genes & repeats, Het: Heterochromatin, PromP: Poised Promoter, PromBiv: Bivalent Promoter, ReprPC: Repressed Polycomb, Qies: Quiescent/Low

表 11: 在胎週数関連エピジェネティックメモリーの候補 CpG 2,093 箇所における 25 種クロマチン状態のエンリッチメント解析結果(ChromHMM における臍帯血 T cell および B cell でのアノテーション (T-, B-cell-Anno) を利用)

*Bonferroni 基準(p -value $<0.05/25$)を満たし、オッズ比 >1 のクロマチン状態に関しては p -value を太字で示した。 **OR: オッズ比、CI: 信頼区間

25-Chromatin-State (T-cell-Anno)	Proportion (%)	OR	95%CI	p -value
Active TSS (TssA)	0.2	0.034	(0.0094, 0.088)	1.3E-42
Promoter upstream TSS (PromU)	8.9	0.82	(0.70, 0.96)	0.010
Promoter downstream with DNase (PromD1)	3.6	0.31	(0.25, 0.40)	4.8E-32
Promoter downstream TSS (PromD2)	2.4	2.5	(1.9, 3.3)	1.4E-08
Transcription 5' (Tx5)	1.0	0.51	(0.31, 0.79)	0.0014
Transcription (Tx)	0.2	0.12	(0.037, 0.27)	7.9E-13
Transcription 3' (Tx3)	1.0	0.18	(0.11, 0.29)	2.7E-24
Weak Transcription (TxWk)	2.1	0.57	(0.41, 0.77)	8.1E-05
Transcription Regulatory (TxReg)	1.2	1.1	(0.70, 1.6)	0.68
Transcription 5' Enhancer (TxEnh5)	0.3	0.53	(0.20, 1.2)	0.13
Transcription 3' Enhancer (TxEnh3)	0.9	1.3	(0.79, 2.1)	0.23
Transcription Weak Enhancer (TxEnhW)	1.0	0.51	(0.31, 0.80)	0.0014
Active Enhancer 1 (EnhA1)	0.9	2.2	(1.3, 3.6)	0.0020
Active Enhancer 2 (EnhA2)	0.9	1.4	(0.81, 2.2)	0.17
Active Enhancer Flank (EnhAF)	0.4	0.83	(0.38, 1.6)	0.76
Weak Enhancer 1 (EnhW1)	5.6	2.8	(2.3, 3.4)	5.9E-21
Weak Enhancer 2 (EnhW2)	4.2	1.3	(1.1, 1.6)	0.012
Enhancer Acetylation Only (EnhAc)	0	0	(0, 3.5)	0.63
DNase only (DNase)	1.5	2.4	(1.6, 3.4)	2.4E-05
ZNF genes & repeats (ZNF_Rpts)	0.2	1.1	(0.29, 2.7)	0.79
Heterochromatin (Het)	0.5	0.76	(0.37, 1.4)	0.49
Poised Promoter (PromP)	3.4	3.1	(2.4, 3.9)	1.4E-15
Bivalent Promoter (PromBiv)	14.0	1.9	(1.7, 2.2)	3.0E-22
Repressed Polycomb (ReprPC)	19.3	3.1	(2.8, 3.5)	3.5E-73
Quiescent/Low (Quies)	26.5	0.79	(0.71, 0.86)	3.1E-07

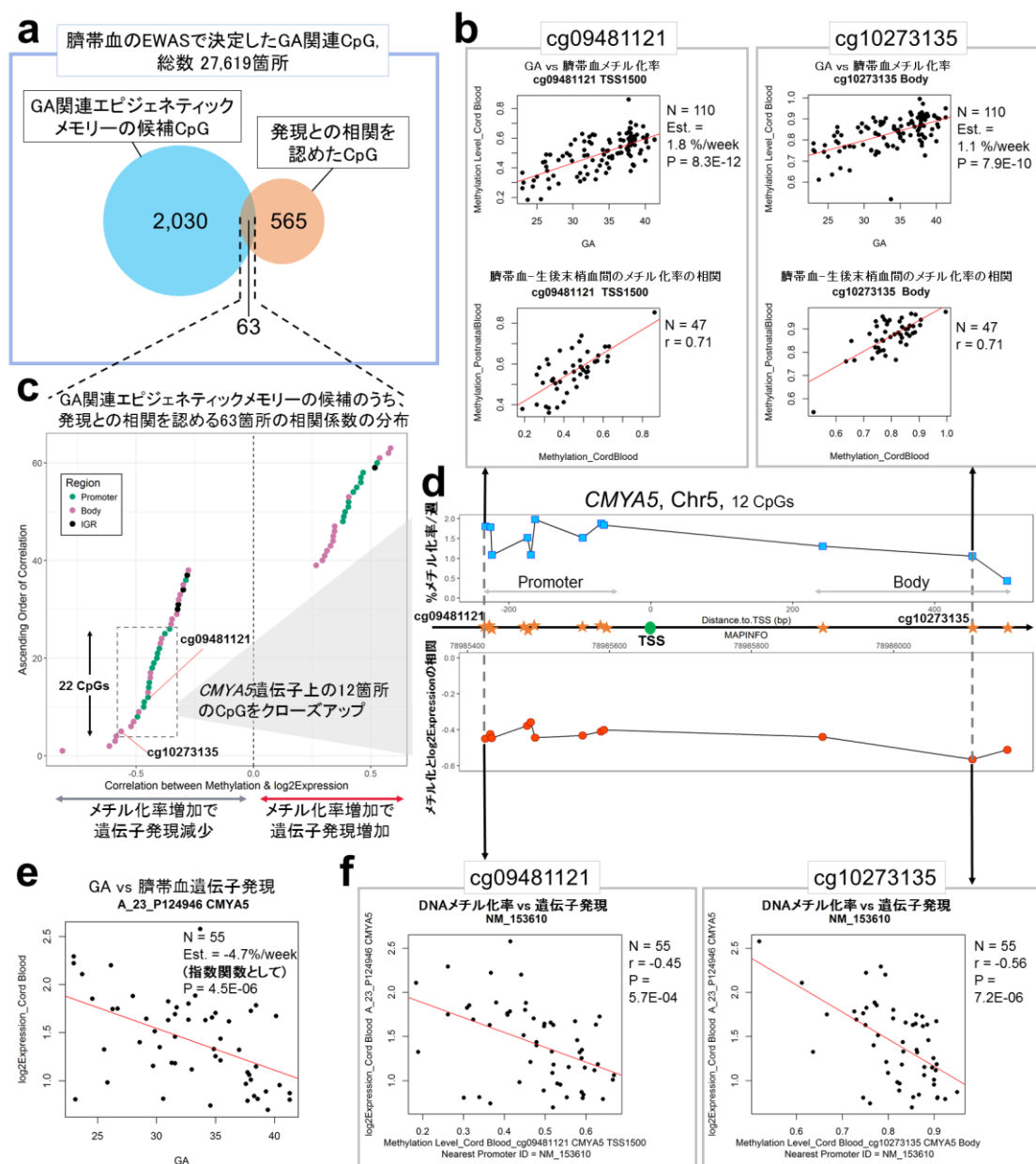
25-Chromatin-State (B-cell-Anno)	Proportion (%)	OR	95%CI	<i>p</i> -value
Active TSS (TssA)	0.6	0.11	(0.056, 0.18)	3.1E-35
Promoter upstream TSS (PromU)	9.7	0.95	(0.82, 1.1)	0.51
Promoter downstream with DNase (PromD1)	6.6	0.48	(0.40, 0.57)	4.2E-20
Promoter downstream TSS (PromD2)	2.1	3.2	(2.3, 4.4)	6.6E-11
Transcription 5' (Tx5)	0.6	0.30	(0.15, 0.52)	4.6E-07
Transcription (Tx)	0.5	0.21	(0.10, 0.39)	2.0E-10
Transcription 3' (Tx3)	1.3	0.21	(0.14, 0.30)	6.4E-27
Weak Transcription (TxWk)	2.0	0.55	(0.40, 0.75)	3.4E-05
Transcription Regulatory (TxReg)	1.6	1.1	(0.78, 1.6)	0.46
Transcription 5' Enhancer (TxEnh5)	0.4	0.52	(0.24, 0.99)	0.051
Transcription 3' Enhancer (TxEnh3)	0.8	1.2	(0.70, 1.9)	0.42
Transcription Weak Enhancer (TxEnhW)	0.7	0.53	(0.29, 0.89)	0.013
Active Enhancer 1 (EnhA1)	0.4	0.68	(0.29, 1.4)	0.37
Active Enhancer 2 (EnhA2)	0.6	0.74	(0.38, 1.3)	0.38
Active Enhancer Flank (EnhAF)	0.4	0.71	(0.33, 1.4)	0.39
Weak Enhancer 1 (EnhW1)	3.9	2.8	(2.2, 3.5)	6.8E-15
Weak Enhancer 2 (EnhW2)	2.9	1.2	(0.92, 1.6)	0.15
Enhancer Acetylation Only (EnhAc)	0.5	1.3	(0.64, 2.3)	0.39
DNase only (DNase)	1.8	2.7	(1.9, 3.8)	8.7E-08
ZNF genes & repeats (ZNF_Rpts)	0.2	0.82	(0.22, 2.1)	1
Heterochromatin (Het)	0.7	0.63	(0.34, 1.1)	0.086
Poised Promoter (PromP)	2.3	2.4	(1.7, 3.1)	1.3E-07
Bivalent Promoter (PromBiv)	14.5	2.3	(2.0, 2.6)	1.3E-33
Repressed Polycomb (ReprPC)	20.8	3.0	(2.7, 3.3)	1.3E-72
Quiescent/Low (Quies)	24.2	0.79	(0.71, 0.87)	1.7E-06

第3位 Bivalent Promoter (転写活性型ヒストンマークの H3K4me3 と転写抑制型ヒストンマークの H3K27me3 の両方の修飾が入る領域; T-cell-Anno: 14.0%, B-cell-Anno: 14.5%)、第4位 Promoter upstream TSS(転写開始点より上流のプロモーター領域; T-cell-Anno: 8.9%, B-cell-Anno: 9.7%)の順である(図 19c、図 21a、表 11)。エンリッチメント解析で、オッズ比が1を超えていてかつ p 値が Bonferroni 基準($p < 0.05/25 = 0.002$)を満たすクロマチン状態はどちらのアノテーションでも Repressed Polycomb, Bivalent Promoter, Poised Promoter(文献により定義が異なるが、本研究で扱う ChromHMM の「25 種クロマチン状態」の分類における定義は、抑制型マーク H3K27me3 との関連の少ない転写活性の弱いあるいは転写活性の不安定なプロモーター領域)、DNase Only(DNA 加水分解酵素である DNase に高感受性なオープンクロマチン領域)、Weak Enhancer 1(活性の弱いエンハンサー領域、ヒストンマークとしては転写活性型の H3K4me1 と H3K4me2 が中心)、Promoter downstream TSS(転写開始点より下流のプロモーター領域、ヒストンマークとしては転写活性型の H3K4me1、H3K4me2、H3K4me3、H3K79me2)であった(図 19d、図 21b、表 11)。

さらに、在胎週数関連エピジェネティックメモリーの候補 CpG サイトにおける遺伝子発現についてこれまでの解析結果を基にまとめ直したので、図 22

に示す。

図 22. 在胎週数関連エピジェネティックメモリーの候補 CpG サイトにおける遺伝子発現



a 水色の円は在胎週数(GA)関連エピジェネティックメモリーの候補 CpG (2,093箇所)、橙色の円は対応する遺伝子に在胎週数と関連する発現を認め(図 17)、DNAメチル化率と遺伝子発現の関連も認めた CpG 628箇所を表す。 **b** CMYA5 上の Promoter 上の CpG サイト cg09481121(左)と Gene Body 上の CpG サイト

cg10273135(右)における、在胎週数と臍帯血 DNA メチル化率の関係を示す散布図(上)と臍帯血-生後末梢血間の DNA メチル化率の関係を示す散布図(下)。 **c** 在胎週数(GA)関連エピジェネティックメモリーの候補 CpG と発現との相関を認めた CpG の交わり部分の 63 箇所における DNA メチル化率-遺伝子発現(log2 スケール)間の相関係数の分布。x 軸が相関係数で、y 軸が相関係数を昇順に並べた際の順位。緑の丸は Promoter 領域(TSS1500, TSS200, 5'UTR, 1stExon)内、ピンクの丸は Gene Body 領域(Body, 3'UTR)内、黒の丸は IGR(intergenic region)内の CpG を意味する。 **d** 水平に走る軸の下に染色体の位置情報(座標)を、上に TSS(転写開始点)に対しての位置関係を示す。星印は 12 箇所の CpG サイトを意味する。上のグラフは各 CpG サイトにおける、在胎週数に対しての DNA メチル化の変化率の推移を示す。下のグラフは各 CpG サイトにおける、DNA メチル化-遺伝子発現(log2 スケール)間の相関係数の推移を示す。 **e** *CMYA5* 遺伝子の発現プロープ A_23_P124946 の在胎週数と遺伝子発現(log2 スケール)の関係を表す散布図。 **f** *CMYA5* 上の Promoter 上の CpG サイト cg09481121(左)と Gene Body 上の CpG サイト cg10273135(右)における、臍帯血 DNA メチル化率と遺伝子発現(log2 スケール)の関係を示す散布図。

*GA: gestational age, EWAS: epigenome-wide association study, Est. : estimates(回帰係数), TSS: Transcription Starting Site(転写開始点)

**TSS200: 転写開始点から上流の 200 塩基までの領域、TSS1500: 転写開始点から上流 200~1500 塩基の間の領域、5'UTR: 5'側の非翻訳領域、3'UTR: 3'側の非翻訳領域

在胎週数関連エピジェネティックメモリー候補 CpG サイト 2,093 箇所の中で、臍帯血サンプルにおいて在胎週数に関連した遺伝子発現を認め、かつ DNA メチル化と遺伝子発現の直接の関連を認めたものは 63 箇所であった(図 22a)。これら 63 箇所のうち DNA メチル化-遺伝子発現の相関が負のもの(メチル化率が増加すると遺伝子発現が減少するもの)は 38 箇所、正のものは 25 箇所であった(図 22b)。

更に TSS1500, TSS200, 5'UTR, 1stExon を Promoter 領域、gene body, 3'UTR を Body 領域として、Promoter、Body、IGR(intergenic region)の 3 つの分類で分けた。

Promoter 領域のメモリー候補は 24 箇所、DNA メチル化-遺伝子発現の相関が負のものは 13 箇所(このうち *CMYA5* 遺伝子内のものが 9 箇所)、正のものは 11 箇所であった。

上記 63 箇所のうち、最も個数の多かった遺伝子が *CMYA5* であり、総数 12 箇所であった。この 12 箇所の内訳は Promoter 領域 9 箇所と Gene Body 領域 3 箇所で、両方の領域の CpG サイトが在胎週数関連の DNA メチル化、在胎週数関連の遺伝子発現、DNA メチル化-遺伝子発現の相関、出生時から生後の維持メチル化を認めた(図 22b, d, e, f)。その他、上記 63 箇所の中で IGR(Intergenic Region: 遺伝子間領域)も含めて、複数箇所の CpG サイトが検出された遺伝子は *UCN*(8 箇所)、*SLC12A7*(5 箇所)、*TNFAIP2*(4 箇所)、*NGF*(3 箇所)、*TAS2R41*(2 箇所)、*EXOC3L4*(Illumina による遺伝子名は *C14orf73*; 2 箇所)、*LPPR2*(2 箇所)であった(表 12: 次頁)。

表 12.

CpGprobe	ChromatinState_Tcell_cord	ChromatinState_Bcell_cord	Gene_Illumina	Gene_Agilent	Correlation	CHR	MAPINFO	feature	%Methylation/week	pval.GA	Cord-Post-Correlation
cg26411822	2_PromU	2_PromU	STX12	STX12	-0.29	1	28099013	TSS1500	-0.7	4.3E-05	0.70
cg00320094	24_ReprPC	23_PromBiv	GPR177	WLS	0.41	1	68695183	Body	0.7	8.4E-08	0.72
cg02987481	24_ReprPC	24_ReprPC	NGF	NGF	-0.41	1	115881544	TSS1500	0.3	5.9E-05	0.73
cg07790079	25_Quies	25_Quies		NGF	-0.32	1	116021925	IGR	0.4	6.5E-05	0.82
cg25599129	25_Quies	25_Quies		NGF	-0.32	1	116022006	IGR	0.7	1.0E-05	0.82
cg16246545	4_PromD2	4_PromD2	PHGDH	PHGDH	-0.52	1	120255941	Body	0.5	1.8E-07	0.70
cg27179622	25_Quies	25_Quies	LEFTY2	LEFTY2	-0.30	1	226127290	Body	0.9	1.3E-04	0.94
cg20028470	24_ReprPC	24_ReprPC	UCN	UCN	0.44	2	27530829	5'UTR	0.3	3.8E-06	0.87
cg07537370	24_ReprPC	24_ReprPC	UCN	UCN	0.39	2	27531124	5'UTR	0.4	4.8E-04	0.79
cg05231308	24_ReprPC	24_ReprPC	UCN	UCN	0.41	2	27531163	TSS200	0.6	2.1E-09	0.84
cg04527918	24_ReprPC	24_ReprPC	UCN	UCN	0.53	2	27531170	TSS200	0.6	4.7E-07	0.82
cg13833437	24_ReprPC	11_TxEnh3'	UCN	UCN	0.46	2	27531236	TSS200	0.6	1.1E-06	0.81
cg05113927	24_ReprPC	11_TxEnh3'	UCN	UCN	0.39	2	27531244	TSS200	0.6	3.7E-06	0.81
cg21752601	24_ReprPC	11_TxEnh3'	UCN	UCN	0.43	2	27531310	TSS200	0.5	4.7E-07	0.72
cg20442078	24_ReprPC	11_TxEnh3'	UCN	UCN	0.46	2	27531360	TSS1500	1.1	4.6E-12	0.73
cg05960677	2_PromU	2_PromU	RBM43	RBM43	-0.33	2	152117363	Body	-1.8	5.0E-09	0.70
cg26627956	6_Tx	6_Tx	CFLAR	CFLAR	0.31	2	202004766	Body	0.7	3.7E-08	0.75
cg21674813	25_Quies	25_Quies	OBSL1	OBSL1	-0.35	2	220425073	Body	0.7	3.9E-08	0.76
cg01948202	3_PromD1	3_PromD1	PARP14	PARP14	-0.32	3	122400474	Body	-1.3	1.3E-07	0.82
cg00509649	9_TxReg	9_TxReg	SLC12A7	SLC12A7	-0.61	5	1103848	Body	-1.3	1.6E-03	0.92
cg13681701	9_TxReg	9_TxReg	SLC12A7	SLC12A7	-0.58	5	1103910	Body	-0.8	5.8E-04	0.90
cg16163535	13_EnhA1	9_TxReg	SLC12A7	SLC12A7	-0.59	5	1107098	Body	-0.8	1.6E-04	0.71
cg14596589	13_EnhA1	9_TxReg	SLC12A7	SLC12A7	-0.44	5	1107148	Body	-0.9	4.1E-04	0.83
cg13592947	3_PromD1	3_PromD1	SLC12A7	SLC12A7	-0.45	5	1111049	Body	-0.6	2.6E-03	0.93
cg09481121	22_PromP	22_PromP	CMYA5	CMYA5	-0.45	5	78985425	TSS1500	1.8	2.1E-13	0.71
cg00611789	22_PromP	22_PromP	CMYA5	CMYA5	-0.42	5	78985432	TSS1500	1.8	4.3E-12	0.81
cg11438310	22_PromP	22_PromP	CMYA5	CMYA5	-0.45	5	78985434	TSS1500	1.1	9.3E-12	0.77
cg10257870	22_PromP	22_PromP	CMYA5	CMYA5	-0.38	5	78985484	TSS200	1.5	5.2E-11	0.75
cg03546977	22_PromP	22_PromP	CMYA5	CMYA5	-0.36	5	78985489	TSS200	1.1	1.7E-10	0.76
cg09655403	22_PromP	22_PromP	CMYA5	CMYA5	-0.44	5	78985495	TSS200	2.0	5.2E-10	0.79
cg15197065	22_PromP	22_PromP	CMYA5	CMYA5	-0.43	5	78985562	TSS200	1.5	2.7E-12	0.75
cg04154027	22_PromP	22_PromP	CMYA5	CMYA5	-0.41	5	78985588	TSS200	1.9	1.9E-12	0.75
cg23279355	22_PromP	22_PromP	CMYA5	CMYA5	-0.40	5	78985592	TSS200	1.8	1.9E-11	0.78
cg21606036	25_Quies	22_PromP	CMYA5	CMYA5	-0.44	5	78985900	Body	1.3	2.9E-12	0.74
cg10273135	25_Quies	25_Quies	CMYA5	CMYA5	-0.56	5	78986111	Body	1.1	7.9E-10	0.71
cg02366931	25_Quies	25_Quies	CMYA5	CMYA5	-0.51	5	78986160	Body	0.4	3.5E-06	0.79
cg11033617	25_Quies	25_Quies	RASGEF1C	RNF130	-0.28	5	179562118	Body	-0.2	1.0E-03	0.81
cg17099072	2_PromU	2_PromU	GABBR1	GABBR1	0.40	6	29601489	TSS1500	0.6	2.0E-05	0.71
cg18221076	17_EnhW2	5_Tx5'		HCG26	0.52	6	31442934	IGR	0.5	9.0E-05	0.83
cg17362900	17_EnhW2	10_TxEnh5'	HLA-DPB1	HLA-DPA1	0.27	6	33047944	Body	0.9	5.1E-08	0.71
cg01418527	17_EnhW2	23_PromBiv	FSCN1	FSCN1	-0.47	7	5631189	TSS1500	0.3	7.8E-04	0.74
cg02557364	25_Quies	25_Quies		TAS2R41	-0.30	7	143208268	IGR	0.2	1.4E-05	0.76
cg08384657	25_Quies	25_Quies		TAS2R41	-0.28	7	143208406	IGR	0.6	1.0E-08	0.73
cg04259752	23_PromBiv	23_PromBiv	LOXL2	LOXL2	0.47	8	23262159	TSS1500	0.6	2.3E-10	0.76
cg02174341	25_Quies	25_Quies	UCMA	UCMA	0.38	10	13276466	TSS200	0.7	2.3E-09	0.71
cg02737384	2_PromU	2_PromU	RASSF4	C10orf10	-0.39	10	45470199	Body	0.6	3.6E-03	0.87
cg12527260	23_PromBiv	2_PromU	RCOR2	RCOR2	-0.40	11	63683744	Body	0.3	8.5E-05	0.80
cg17205324	24_ReprPC	24_ReprPC	EFS	EFS	-0.47	14	23835595	TSS1500	0.4	5.7E-12	0.80
cg19628497	24_ReprPC	24_ReprPC	DLK1	DLK1	-0.81	14	101194267	Body	0.8	1.7E-03	0.93
cg22502625	24_ReprPC	24_ReprPC	C14orf73	EXOC3L4	-0.31	14	103568366	Body	0.3	1.8E-03	0.78
cg10973146	24_ReprPC	24_ReprPC	C14orf73	EXOC3L4	-0.35	14	103568471	Body	0.7	7.9E-07	0.72
cg09814127	5_Tx5'	8_TxWk	TNFAIP2	TNFAIP2	0.35	14	103593235	Body	0.5	3.1E-04	0.89
cg18587137	5_Tx5'	8_TxWk	TNFAIP2	TNFAIP2	0.30	14	103593503	Body	0.9	1.9E-05	0.87
cg18620571	5_Tx5'	8_TxWk	TNFAIP2	TNFAIP2	0.34	14	103593505	Body	0.7	1.8E-04	0.91
cg10501093	5_Tx5'	8_TxWk	TNFAIP2	TNFAIP2	0.33	14	103593520	Body	0.8	1.8E-04	0.89
cg05899103	8_TxWk	8_TxWk	LOC390595	UBAP1L	0.58	15	65394570	Body	0.7	9.7E-09	0.72
cg24339470	16_EnhW1	23_PromBiv	SARM1	SARM1	0.34	17	26712412	Body	0.2	3.3E-05	0.72
cg08577293	19_DNase	19_DNase	DHX58	DHX58	0.54	17	40253614	3'UTR	0.7	1.9E-07	0.72
cg00441918	19_DNase	19_DNase	IGF2BP1	IGF2BP1	-0.49	17	47113407	Body	0.5	7.1E-05	0.78
cg00376735	8_TxWk	7_Tx3'	LPPR2	LPPR2	0.35	19	11473212	Body	1.1	5.0E-10	0.73
cg19699893	8_TxWk	7_Tx3'	LPPR2	LPPR2	0.29	19	11473353	Body	0.4	7.2E-08	0.71
cg17250863	8_TxWk	8_TxWk	GGT7	GGT7	0.58	20	33451272	Body	1.0	8.3E-12	0.75
cg23835646	2_PromU	13_EnhA1	LGALS2	LGALS2	-0.49	22	37976063	TSS200	-0.4	4.6E-05	0.73

(表 12 説明) 在胎週数関連エピジェネティックメモリーの候補 CpG の中で発現との相関を認めた CpG63 箇所のクロマチン状態および DNA メチル化変化、遺伝子発現との相関、生後のメチル化維持に関するまとめ

CpGprobe: DNA メチル化アレイにおけるプローブの ID

ChromatinState_Tcell_Cord: 臍帯血 T cell でのクロマチン状態のアノテーション。色分けは ChromHMM に準じた。

ChromatinState_Bcell_Cord: 臍帯血 B cell でのクロマチン状態のアノテーション

Gene_Illumina: イルミナのアノテーションによる対応遺伝子の名前。2 つ以上同一の遺伝子名を認める場合はクリーム色で示した。

Gene_Agilent: アジレントのアノテーションによる対応遺伝子の名前。2 つ以上同一の遺伝子名を認める場合はクリーム色で示した。

Correlation: DNA メチル化と遺伝子発現(log2 スケール)の間の相関係数。負の相関は青で、正の相関は赤で表した。

CHR: 染色体の番号

MAPINFO: 染色体上の位置情報(座標)

feature: TSS を中心に分類した遺伝子上の領域名。図 8, 11, 16 の説明を参照。

Promoter 領域(TSS1500, TSS200, 5'UTR, 1stExon)は緑色、Gene Body 領域(Body, 3'UTR)はピンク色、IGR(intergenic region)は灰色で表した。

%Methylation/week: 在胎週数についての EWAS で「モデル 1」で算出された回帰係数。1 週間あたりのメチル化率の変化(%)を表す。

pval.GA: 在胎週数についての EWAS で「モデル 1」で算出された *p*-value。

Cord-Post-Correlation: 臍帯血-生後末梢血間のメチル化率の相関係数。相関係数の高いものを濃い赤で表した。

EWAS: epigenome-wide association study

本項の最後に、確認・参考のため、エピジェネティックメモリーの候補絞り込みに臍帯血・生後児血の DNA メチル化データを用いた児 47 人の臨床データの分布を以下、表 13 に示す。平均の在胎週数は 30.9 週で平均の SD スコアは-1.0 であった。44.7%が SGA(ここでは出生体重が 10 パーセンタイル未満)で、6.4%

が large for gestational age (LGA; ここでは出生体重が 90 パーセントイル以上)であった。48.9%の児が男児であった。83.0%の児が帝王切開で出生していた。母体の平均年齢は 34.0 歳であった。母体の平均妊娠前 BMI は 22.1 であった。妊娠中の母体喫煙は 1 人(2.1%)であった。4 人の母体(8.5%)が妊娠前まで喫煙していた。6.4%(3 人)の母体が妊娠糖尿病、31.9%が絨毛膜羊膜炎、14.9%が絨毛膜羊膜炎などの炎症を伴わない特発性の前期破水、29.8%が妊娠高血圧症を合併していた。

表 13. エピジェネティックメモリーの候補絞り込みの解析に臍帯血・生後児血の DNA メチル化データを用いた 47 人の児の臨床データの分布

	平均値 (SD)	中央値	N (%)
在胎週数	30.9 (3.5)		31.6
<28 週			11 (23.4)
28-32 週			14 (29.8)
32-37 週			21 (44.7)
>37 週			1 (2.1)
出生体重 SD スコア	-1.0 (1.5)		-0.9
<-2.5			11 (23.4)
-2.5 ~ -1.28			10 (21.3)
-1.28 ~ 1.28			23 (48.9)
1.28 ~ 2.5			3 (6.4)
性別			
男児			23 (48.9)
女児			24 (51.1)

	平均値 (SD)	中央値	N (%)
分娩様式			
帝王切開			39 (83.0)
経膣			8 (17.0)
母体妊娠前 BMI	22.1 (4.3)	21.5	
<18.5 kg/m ²			10 (21.3)
18.5 ~ 25 kg/m ²			30 (63.8)
25 ~ 30 kg/m ²			2 (4.3)
>30 kg/m ²			5 (10.6)
妊娠中の母体喫煙			
Yes			1 (2.1)
No			46 (97.9)
妊娠前の母体喫煙			
Yes			4 (8.5)
No			43 (91.5)
絨毛膜羊膜炎 (CAM)			
Yes			15 (31.9)
No			32 (68.1)
特発性の前期破水 (iPROM)			
Yes			7 (14.9)
No			40 (85.1)
妊娠高血圧症			
Yes			14 (29.8)
No			33 (70.2)
母体年齢	34.0 (4.9)		34
<25 歳			0 (0)
25 ~ 30 歳			8 (17.0)
30 ~ 35 歳			16 (34.0)
35 ~ 40 歳			18 (38.3)
>40 歳			5 (10.6)

	平均値 (SD)	中央値	N (%)
父親の年齢	37.4 (7.7)		36
<25 歳			1 (2.1)
25 ~ 30 歳			4 (8.5)
30 ~ 35 歳			11 (23.4)
35 ~ 40 歳			14 (29.8)
>40 歳			17 (36.2)
父親の BMI	23.4 (3)		23.1
<18.5 kg/m ²			1 (2.1)
18.5 ~ 25 kg/m ²			35 (74.5)
25 ~ 30 kg/m ²			9 (19.1)
>30 kg/m ²			2 (4.3)
生殖補助医療による妊娠 (ART)			
Yes			11 (23.4)
No			36 (76.6)
母体の妊娠糖尿病 (GDM)			
Yes			3 (6.4)
No			44 (93.6)
前置胎盤			
Yes			6 (12.8)
No			41 (87.2)
経産数			
>0			20 (42.6)
0			27 (57.4)

IV. 考察

低出生体重児は早産児・Small for Gestational Age(SGA)児の2つに大きく分類されるが、いずれも2型糖尿病や冠動脈疾患、神経学的問題などのハイリスク児である。「子宮内・生後早期における低酸素・低栄養など悪環境への適応に伴い、疾病体質が形成される」というのがDOHaD仮説であるが、エピジェネティックなメカニズムの関与がこれまでに示唆されてきた。本研究の目的は臍帯血および生後末梢血中の単核球分離サンプルにおけるDNAメチル化および遺伝子発現データを用いて、早産児およびSGA児に起こるエピジェネティック変化を調べ、そのうち後生に持続する候補を探索することである。本研究は早産およびSGA児の分野において、(1)日本における、血液検体でのエピゲノムワイド関連解析(EWAS)を行った最初の研究であり、(2)同一の単核球分離後サンプル内で発現も広範囲に検討した世界で初めてのEWAS研究である。

本研究では最初に早産やSGAに影響しうる臨床的特徴を検討したが、絨毛膜羊膜炎[58]、特発性の前期破水[59]、妊娠高血圧症[60]が在胎週数と負の相関すなわち早産との関連を有し、妊娠高血圧症が出生体重SDスコアと負の相関すなわちSGAとの相関を有して[60]いたことは先行疫学研究と合致する所見であった。二番目に必須の調整変数である細胞組成、実験バッチ、性別に前述の

臨床的特徴や妊娠前喫煙、分娩様式(帝王切開)、妊娠前母体 BMI を加えて、線形回帰モデルによる調整を行い在胎週数ないし出生体重 SD スコアについての EWAS を行った。150 箇所 of SD スコア関連 CpG サイトと 27,619 箇所の非常に多くの在胎週数関連 CpG サイトを検出した。非常に多くの在胎週数関連 CpG サイトが検出されたことは先行の EWAS 研究[24, 25]でも述べられており、また CpG の所在をみても CpG アイランドを軸とした分類では CpG アイランドの隣の CpG shore が最多数だったことも先行研究[24]と合致する。先行研究では在胎週数が進むとメチル化率が下がるメチル化率減少型在胎週数関連 CpG サイトの方が多く検出されていた[21, 22, 25]のに対し、本研究では在胎週数が進むとメチル化率が増加する CpG サイトが多く検出された。これは、先行研究が全て好中球を多く含有する臍帯血全血ないし全血に準じたサンプルを用いた研究なのに対し、本研究はリンパ球がメインの単核球分離後サンプルを用いていることに起因するのかもしれない。ゲノム全般的な傾向として、成熟により顆粒球は低メチル化方向に向かい、リンパ球は高メチル化方向に進むことが最近の基礎的な研究から推察される[47]。

その後、上述の EWAS で得られた在胎週数関連 CpG サイトおよび出生体重 SD スコア関連 CpG サイトを併せた全体の約 7 割程度にあたる 19,640 箇所の

CpG サイトと対応する遺伝子転写産物 9,691 箇所を焦点を当てて、遺伝子発現関連解析を行なった。この解析では多重検定補正後有意基準(FDR 基準)を満たす SD スコア関連遺伝子発現は検出されず補正なしの場合でも 6 箇所と非常に少なかったが、401 種(RefSeq 遺伝子としては 369 個)と多くの在胎週数関連遺伝子発現(FDR 基準)を認め、補正なしの場合は 1,834 個に上った。在胎週数関連遺伝子発現が多数確認された一方で SD スコア関連遺伝子発現が検出されなかったのは、サンプル数が 55 と少なかったことが大きな理由の一つかもしれないが、関連 CpG サイトの所在として在胎週数関連 CpG はより発現に影響しやすい CpG shore に多く存在し、SD スコア関連 CpG は発現に影響しにくい open sea に多く存在していたこととも関係しているのかもしれない。また、興味深いことに、補正なしの在胎週数関連発現遺伝子名を用いたパスウェイ解析では在胎週数が進むと発現が減少する発現減少型遺伝子発現のカテゴリーでは PI3K-Akt signaling pathway と ECM-receptor interaction が、発現増加型のカテゴリーでは NF-kappa B signaling pathway が有意かつ上位に位置したが、これはメチル化解析で得られた在胎週数関連プロモーター領域内 CpG サイト責任遺伝子を用いたパスウェイ解析の結果と矛盾しない結果である。すなわち、EWAS 後パスウェイ解析では、在胎週数が進むとメチル化が増加するメチル化率増加型のカテゴリ

ーで PI3K-Akt signaling pathway と ECM-receptor interaction が、メチル化率減少型のカテゴリーで NF-kappa B signaling pathway が上位だったのだが、これは「DNA メチル化が増加すると遺伝子発現が抑制される」定説に従う結果となり、実験や解析の妥当性を支持する結果と考えられる。また、PI3K-Akt signaling pathway は糖代謝・神経発達・サイトカイン・細胞増殖などに関与する重要なパスウェイの一つであり、そのような意味でも興味深い。DNA メチル化と遺伝子発現の関連・相関についても確認を行ったが、関連が見られたのが 369 個の RefSeq 遺伝子中約 3 分の 2 の 257 個であった。このうち負の相関、すなわちメチル化率が増加すると発現が減少する、定説型の相関をとるものは約 3 分の 2 で、残りの 3 分の 1 は正の相関、反定説型の相関であった。しかし、最近の研究結果から、反定説型のメチル化-発現関係は少なからず見られることが分かっており [47, 61]、割合としても他の分野の先行研究と大きな差はなかった[61]。

最後の解析として、出生時の DNA メチル化変化が生後も持続するエピジェネティックメモリーの候補を選ぶ解析を行った。在胎週数関連 CpG サイト 27,619 箇所のうち、2,093 箇所が候補として選択されたが、例として挙げた *GABBR1* 遺伝子の CpG は場所の完全一致はないものの、同一遺伝子内の他の CpG ではエピジェネティックメモリーを示唆する報告がある[30]。 *GABBR1* は

GABA-B 受容体 1 の責任遺伝子で神経細胞においては抑制性のシグナル伝達を担う[62]。本研究で用いた細胞は、リンパ球系を主とする血液細胞であり、神経細胞に同様のエピジェネティック変化が起こるとは言えないが、解析上は PI3K-Akt signaling pathway など GABA に関連するパスウェイも検出されており、パスウェイという意味で何らか他の臓器でも起こっている可能性を考慮して、代表的な機能の一つをここで紹介した。TRHR 遺伝子は甲状腺ホルモン放出ホルモン受容体の責任遺伝子だが、早産児ではしばしば甲状腺ホルモンが低値になり、治療対象となることもあり[63]、興味深かったので、例示した。CACNA1C は電位依存性カルシウムチャンネルサブユニット α_{1c} の責任遺伝子でこの遺伝子のメチル化状態が双極性障害との関連を示唆する報告[64]もある。また、マウスではこのカルシウムチャンネルの機能低下が心肥大や心不全をもたらすとする報告もある[65]。PRDM16 は近年の研究で褐色脂肪組織の発達に関わり、肥満との関わりが注目されている遺伝子である[66]。また、これら 2,093 箇所のエピジェネティックメモリーの候補についてクロマチン状態の特徴がないかどうかを ChromHMM の 25-Chromatin-State のアノテーションを用いてエンリッチメント解析を行ったが、Repressed Polycomb(抑制性のポリコーム結合部位)と Bivalent Promoter(転写活性型ヒストンマークの H3K4me3 と転写抑制型ヒストンマーク

の H3K27me3 の両方の修飾が入る Promoter 領域)が数的にも多く、また有意レベルも非常に高く特にエンリッチされていた。これら 2 つのクロマチン状態は 2 者の中で移行しやすく類似のクロマチン状態ともいえる。

上述のエピジェネティックメモリー候補の CpG サイト 2,093 箇所のうち、臍帯血で DNA メチル化と遺伝子発現との相関が確認できたのは、僅かに 63 箇所であるが、プロモーター領域で反定説型発現パターンをとるものはさらに無視できない割合となった。また、Gene Body 領域はメチル化率が増加すると発現が促進されることが多い[29]と考えられているが、それに反するものも無視できないレベルで一定数存在することが分かった。例えば、Urocortin の責任遺伝子 UCN 上の CpG サイト (cg200228470, cg07537370, cg05231308, cg04527918, cg13833437, cg05113927, cg21752601, cg20442078) は全て Promoter 領域内に位置しているが、発現とは正の相関(メチル化率増加に伴い発現増加)を示す。これら UCN 上の CpG サイトのクロマチン状態(25-Chromatin-state)は、半数は T-cell でのアノテーション(T-cell-Anno)、B-cell でのアノテーション(B-cell-Anno)の両方、半数は T-cell-Anno で Repressed Polycomb(抑制性のポリコム結合部位)であった。また、C14orf73(EXOC3L4)遺伝子上の CpG サイト (cg22502625, cg10973146) は Gene Body 内に存在するが、発現とは負の相関(メチル化率増加に伴い発現減少)

を示したが、クロマチン状態は T-cell-Anno, B-cell-Anno の両方で Repressed Polycomb であった。最近の研究で、ポリコーム結合部位でのメチル化変化で定説とは異なる発現パターンをとることが報告[67, 68]されており、これらの CpG サイトの発現パターンに関しても同様にポリコームの影響が関与しているのかもしれない。

エピジェネティックメモリーの候補 CpG サイトで遺伝子発現との関連がみられた 63 箇所の中で、3 箇所以上が同一遺伝子内にあったものについて CpG サイトの数が多し順に述べる。CMYA5 遺伝子内には Promoter 領域に 9 箇所の CpG サイトと Gene Body 領域に 3 箇所、計 12 の CpG サイトが検出された。CMYA5 遺伝子は主に骨格筋や心筋に発現している遺伝子であり[69]、変異や遺伝子多型により心室肥大・血圧上昇[70, 71]や、統合失調症の発症リスクが上昇する[72]ことも報告されている遺伝子であるが、Promoter 領域内のクロマチン状態は T-cell-Anno、B-cell-Anno とともに Poised Promoter(文献により定義が異なるが、本研究で扱う ChromHMM の「25 種クロマチン状態」の分類における定義は、抑制型マーク H3K27me3 との関連の少ない転写活性の弱いあるいは転写活性の不安定なプロモーター領域)であった。在胎週数関連のメチル化および発現、メチル化と発現の負の相関(メチル化率増加に伴い、発現減少)がみられたが、Promoter

領域よりも Gene Body 領域の方が強い相関がみられた。しかし、Promoter 領域の方が CpG サイトの数としては多く、数的な要素が実際の発現に寄与したものと考えた。Urocortin の責任遺伝子 *UCN* 上の CpG サイトは上述の通り、プロモーター領域内に 8 箇所の CpG サイトを認め、メチル化-発現関係は反定説型の正の相関を示していたが、ポリコームによる影響が示唆された。Urocortin は Corticotropin-releasing-hormone(CRH: 副腎皮質刺激ホルモン放出ホルモン)の family 中の 1 分子で、神経細胞内で多く認められ、心血管系・消化管系・免疫系などのストレス反応の際に重要な役割を担うものと考えられている[73]。中でも血管拡張作用や心筋虚血後再灌流による傷害に対しての心筋保護作用が報告されている[73, 74]。*SLC12A7* 遺伝子内には、5'から 3'にかけて Promoter 領域・Gene Body 領域と定義されたイルミナのアノテーションにおける Gene Body 領域に 5 箇所の CpG サイトを認め、これらはメチル化-発現状態は負の相関である。しかし、クロマチン状態としては他の 4 箇所(cg00509649, cg13681701, cg16163535, cg14596589)は Transcription Regulatory(転写制御部位)、Active Enhancer 1 (活動型エンハンサー1)などのアノテーションがついているものの、最も 3'側の cg13592947 は Promoter downstream with DNase(DNase の高感受性を有する下流側のプロモーター)のアノテーションがついており、実際には下流型

のプロモーターであるため、通常の Body 領域での発現パターンではなくなる。

SLC12A7 は別名 *KCC4* とも呼ばれ、Potassium-chloride cotransporter(カリウム-クロライド共輸送体)の中の一つであり、神経細胞など多くの細胞にみられ[75]、神経伝達において重要な役割を果たし[76]、GABA の分泌への関与も報告されている[77]が、その他腫瘍細胞の浸潤に影響することも報告[78, 79]されている。

TNFAIP2 遺伝子内には 4 箇所の CpG サイトが該当し、全て Body 領域内だが、メチル化-発現関係は正の相関を示し、一般的な発現パターンであった。*TNFAIP2* 遺伝子は TNF- α induced protein 2 の責任遺伝子で、*EXOC3L3* 遺伝子とも呼ばれ、前述の抑制性のポリコームとの関連について述べた *EXOC3L4* とは 24,764 塩基(約 25kb)と比較的近い位置に存在する。両遺伝子のメチル化とも、在胎週数が進むにつれ増加する傾向にあるが、*EXOC3L4* の発現は減少し、*TNFAIP2* の発現は増加する。この遺伝子は、NF- κ B の標的遺伝子の一つ[80]であり、TLR(Toll-like receptor) signaling など免疫系に関与[81]し、腫瘍[82, 83]や自己免疫疾患[81, 84]との関連が報告されている。*NGF* 遺伝子は近傍の Intergenic Region(IGR: 遺伝子間領域)も含めて、3 箇所(CG02987481: TSS1500, CG07790079 & CG25599129: IGR)の CpG サイトが該当した。Promoter 領域と意義は同じものと考えられ、メチル化-発現関係は負の相関を示し、定説に合致するパターンであった。*NGF* 遺伝子

は NGF(Nerve Growth Factor)の責任遺伝子で、NF-kappa B signaling pathway の構成要素の一つである。神経伝達・軸索輸送に関与する遺伝子で、変異により自律神経障害が起こる[85-87]。近年ではアルツハイマー病における遺伝子治療の標的として注目を集めている[88, 89]。

再度 2,093 箇所のエピジェネティックメモリーの候補 CpG サイトについてまとめ直すと、これらの箇所には一定のクロマチン状態の傾向がみられ、大半は発現との直接の相関がなかった。また少数の発現との関連がみられた箇所には、反定説型の発現パターンをとるものが少なからずみられ、ポリコム修飾の好発部位も一定頻度でみられた。エピジェネティックメモリーの形成には DNA メチル化そのものとは別のエピジェネティックなメカニズムが関与していることが推察された。

研究の限界として、やはりサンプル数が最大でも 110 と限られていることが第一に挙がる。先行研究[21-23, 26, 30, 36]で用いられている全血もしくは全血に準じたサンプルに比べ、新生児血からの単核球分離は実験工程において労力が要求されるため、結果としてサンプル数が減ったといえるかもしれない。しかし、この方法の選択で、細胞集団がリンパ球集団により絞られた効果もあったと考えている。最近の基礎的な研究[47]でも示されているが、骨髄球は成熟

により低メチル化に向かい、リンパ球は成熟に伴い高メチル化に向かう、という真逆の性質を持つ。先行研究のように全血もしくは全血に準じたサンプルを用いた場合は、このように真逆の性質を持つ 2 大系統の細胞群が均等なレベルで内在することになり、DNA メチル化と遺伝子発現ともに評価が不正確になる可能性がある。単核球分離を行なった場合は、リンパ球系統の DNA メチル化と遺伝子発現が反映されることが予想される。「血球の中でどの細胞分画の DNA メチル化が DNA メチル化によるエピジェネティック変化の全身への影響を一番的確に反映していると考えられるのか」、というのは分かっておらず、好中球など顆粒球に絞り込む方法も選択できたが、現実的に好中球よりもリンパ球の方が半減期も長く、好中球分離よりも単核球分離の方が簡便で実験手技上の利点があると考えたので、単核球分離を行なった。本研究では、リンパ球の中で最大数をとることの多い CD4T cell 由来の DNA と RNA が、メチル化と遺伝子発現ともに結果に最も寄与したものと考えている。このようにして、DNA メチル化と遺伝子発現について同時かつより正確に評価することが可能となり、これまでの全血サンプルの DNA メチル化のみの評価に留まる研究で得られなかった知見が集められたことや、骨髄球と異なるリンパ球が主体のサンプルの特性を新たに報告できたことは本研究の強みである。第 2 の限界として、2 施設の

みで得られたサンプルを扱っており、また解析によりサンプル数および集団属性が異なる(臍帯血の EWAS: n = 110、臍帯血の遺伝子発現関連解析およびメチル化-発現間の関連解析: n = 55、臍帯血・生後児血の DNA メチル化データを用いたエピジェネティックメモリーの候補絞り込みの解析: n = 47)ため、集団バイアスは懸念される。特に最後のエピジェネティックメモリー候補絞り込みの解析では、ほとんどが早産児のサンプルからのデータからによるものとなった。しかし、少なくとも臍帯血の EWAS に関しては、限られた集団の中においても早産児の一般的な傾向も確認され、解析手法に関しては適切に対処したものと考えている。また、エピジェネティックメモリーの候補でメチル化-発現間の関連もみられた *CMYA5*, *UCN*, *SLC12A7* 上の大部分の CpG サイトは最近の先行 EWAS 研究[25]でも在胎週数関連 CpG として確認されている。更に、エピジェネティックメモリーの候補 CpG にはポリコーム結合の好発部位のエンリッチメントを確認しており、生物種は異なるが最近の基礎的な研究でポリコームがエピジェネティックメモリーの形成で重要な役割を担うといわれている[90, 91]。これなどを考慮すると、最後の解析についても一定の妥当性はあるものと考えられる。今後、サンプル数や施設数を増やしての日本人集団での検討が望まれる。第 3 の限界として、エピジェネティックメモリーの評価時期が予定日付近のみにな

ってしまったことが挙げられる。DOHaD 仮説は「胎児期ないし幼少期の環境曝露による成人後疾病体質の形成」を基本概念としており、エピジェネティックメモリーを評価するための生後検体採取時期として、最も理想的なのは成人期や学童期と考えるのが自然である。実際に、縦断的に評価している先行研究もあり、先行研究[23]では7歳時と17歳時に、先行研究[30]では18歳時に生後検体を採取・評価している。[23]では7歳以降にほぼすべての在胎週数関連メチル化変化がなくなり、[30]ではわずかに在胎週数関連メチル化変化が残っている。この2つの研究の結論は食い違う点もあるが、大部分の在胎週数関連メチル化変化がなくなったという点では共通している面もある。本研究としては、我が国での母子コホート設立の歴史は日が浅く、データベースやフォローアップ体制も確立されていない現状もあり、児の退院後に生後検体を採取することが極めて困難であるために、生後1年以降の採取は行なわなかった。しかし、生後早期にどのような変動をとるかに関しては先行研究ではまとまった評価がされず、そのような意味で修正40週前後でのDNAメチル化評価をしたことは逆に本研究の強みである。また、先行研究[23, 30]の学童期・成人期近くのEWASでは大部分のメチル化変化がなくなっており、メチル化変化として残りやすい部分を見つける意味では、言い換えるとエピジェネティックメモリーの

根拠は本当に存在するのかを判定する意味では、むしろこの早めの時期の生後検体採取・評価に特有の強みがあるのではないかと考えている。第 4 の限界として、あくまでリンパ球を主とする血球の評価に留まったことが挙げられる。

DOHaD の前身である Barker 仮説では、2 型糖尿病やメタボリックシンドロームに焦点を当てられており、可能であれば肝臓・膵臓をはじめとする臓器のエピジェネティックな状態を評価したいと考えるのが自然であるが、実際には臓器の検体採取には制限・制約がある。血球は最もアクセスしやすい細胞の一つで、その DNA メチル化に代替的マーカーとしての意味を期待することが多い反面、他臓器への応用には注意を要するという意見[92]もある。そもそも、細胞・臓器が異なれば、発生・分化の過程が異なるのでエピジェネティックには異なる。

このような批判的な意見も踏まえた上で、今回の解析結果の細胞・組織・臓器としての意義を述べると、まず DOHaD の概念は発祥となった Barker 仮説のように代謝系などに焦点を当てたものではなく、包括的なもので疾病全般に適用されるので、純粋に血液・免疫系における一定の評価にはなると考える。そして、前述の反対意見[92]もある一方で、他の分野では特定の遺伝子領域では血球と脂肪組織との類似性も見られた報告[93, 94]もあり、今後のデータの蓄積と整理により正しい遺伝子領域選択ができれば他の組織・臓器への意義づけも可能

になるものと期待される。第 5 の限界として、周産期に「結果」として生じた DNA メチル化変化の評価に留まったことが挙げられる。すなわち、DNA メチル化は「結果」と「原因」の双方の性質を併せ持つ。今後は周産期に発生した DNA メチル化変化からの将来の疾病予測が重要になってくると考えられるが、そのためにはコホートやバイオバンクとの連携が必要不可欠と考える。

V. 結論

エピゲノムワイド関連解析(EWAS)により、多くの在胎週数関連 DNA メチル化変化、およびいくつかの出生体重 SD スコア関連メチル化変化を確認し、前者においては遺伝子発現との関連も一定数確認した。在胎週数関連メチル化変化の中には生後にも持続する候補を見いだすことができ、早産児に在胎週数関連のエピジェネティックメモリーが発生することが示唆された。これらのエピジェネティックメモリーはポリコーム蛋白の結合などクロマチン状態の傾向が存在することも分かり、また必ずしも遺伝子発現との関連が確認できない場合や、DNA メチル化増加に伴い遺伝子発現が増加するなど定説にそぐわない場合も一定数見られることが分かり、エピジェネティックメモリーの形成にはメチル化そのものだけではなく、メチル化以外のエピジェネティックなメカニズムが関与していると考えられた。本研究は今後の日本の新生児 EWAS 研究の基盤になるものとする。

倫理面への配慮

本研究は、東京大学医学部附属病院ヒトゲノム倫理委員会(G10036)、国立成育医療研究センター倫理委員会(234)、東京都立墨東病院倫理・個人情報保護委員会(38)の承認を得て行った。母児の検体および臨床情報の提供に関しては全て、面談で書面および口頭で説明した後、書面での同意を得ている。

謝辞

本研究は国立研究開発法人日本医療研究開発機構（AMED）の成育疾患克服等次世代育成基盤研究事業の支援によって行われた。

ご指導いただきました東京大学医学部附属病院 小児・新生児集中治療部の高橋尚人教授、国立成育医療研究センター研究所 周産期病態研究部の秦健一郎部長、河合智子室長、中林一彦室長、嘉村浩美氏、同メディカルゲノムセンターの青砥早希氏、同分子内分泌研究部の松原圭子氏、同社会医学研究部の浦山ケビン部長、同感染免疫・アレルギー研究部の松本健治部長、武田和江氏、東京大学医学部附属病院小児科の西村力助教、岩手医科大学いわて東北メディカル・メガバンク機構の清水厚志教授、志波優講師、東京大学大学院医学

系研究科国際保健学専攻国際生物医科学講座発達医科学分野の水口雅教授、同
生殖・発達・加齢医学専攻小児医学講座の岡明教授、同生殖・発達・加齢医学
専攻産婦人科学講座の藤井知行教授、永松健准教授、東京都立墨東病院産婦人
科の久具宏司部長、兵藤博信部長、同新生児科の清水光政部長、大森意索部長
に深謝申し上げます。また、検体処理にご協力いただいた東京大学医学部附属
病院産婦人科および小児科の医師・看護師の方々、東京都立墨東病院産婦人科
および新生児科の医師・看護師の方々に深謝申し上げます。

参考文献

1. Barker, D.J. and C. Osmond, *Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales*. Lancet. **1**(8489): p. 1077-81 (1986).
2. Barker, D.J., *In utero programming of chronic disease*. Clin Sci (Lond). **95**(2): p. 115-28 (1998).
3. Aarnoudse-Moens, C.S., et al., *Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children*. Pediatrics. **124**(2): p. 717-28 (2009).
4. Roseboom, T.J., et al., *Effects of prenatal exposure to the Dutch famine on adult disease in later life: an overview*. Mol Cell Endocrinol. **185**(1-2): p. 93-8 (2001).
5. Bhargava, S.K., et al., *Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood*. N Engl J Med. **350**(9): p. 865-75 (2004).
6. Hofman, P.L., et al., *Premature birth and later insulin resistance*. N Engl J Med. **351**(21): p. 2179-86 (2004).
7. Irving, R.J., et al., *Adult cardiovascular risk factors in premature babies*. Lancet. **355**(9221): p. 2135-6 (2000).
8. Gluckman, P.D. and M.A. Hanson, *Developmental origins of disease paradigm: a mechanistic and evolutionary perspective*. Pediatr Res. **56**(3): p. 311-7 (2004).
9. Gluckman, P.D., et al., *Effect of in utero and early-life conditions on adult health and disease*. N Engl J Med. **359**(1): p. 61-73 (2008).
10. Bateson, P., et al., *Developmental plasticity and human health*. Nature. **430**(6998): p. 419-21 (2004).
11. Bousquet, J., et al., *Developmental determinants in non-communicable chronic diseases and ageing*. Thorax. **70**(6): p. 595-7 (2015).
12. Chen, T., et al., *Developmental origins of inflammatory and immune diseases*. Mol Hum Reprod. **22**(8): p. 858-65 (2016).
13. Krishna, M., et al., *Birth size, risk factors across life and cognition in late life: protocol of prospective longitudinal follow-up of the MYNAH (MYsore studies of Natal effects on Ageing and Health) cohort*. BMJ Open. **7**(2): p. e012552 (2017).

14. Walker, C.L. and S.M. Ho, *Developmental reprogramming of cancer susceptibility*. Nat Rev Cancer. **12**(7): p. 479-86 (2012).
15. Loh, K.M. and B. Lim, *Epigenetics: Actors in the cell reprogramming drama*. Nature. **488**(7413): p. 599-600 (2012).
16. Rajagopal, J. and B.Z. Stanger, *Plasticity in the Adult: How Should the Waddington Diagram Be Applied to Regenerating Tissues?* Dev Cell. **36**(2): p. 133-7 (2016).
17. *Moving AHEAD with an international human epigenome project*. Nature. **454**(7205): p. 711-5 (2008).
18. Rakyan, V.K., et al., *Epigenome-wide association studies for common human diseases*. Nat Rev Genet. **12**(8): p. 529-41 (2011).
19. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types*. Nature. **473**(7345): p. 43-9 (2011).
20. Ernst, J. and M. Kellis, *Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues*. Nat Biotechnol. **33**(4): p. 364-76 (2015).
21. Schroeder, J.W., et al., *Neonatal DNA methylation patterns associate with gestational age*. Epigenetics. **6**(12): p. 1498-504 (2011).
22. Parets, S.E., et al., *Fetal DNA Methylation Associates with Early Spontaneous Preterm Birth and Gestational Age*. PLoS One. **8**(6): p. e67489 (2013).
23. Simpkin, A.J., et al., *Longitudinal analysis of DNA methylation associated with birth weight and gestational age*. Hum Mol Genet. **24**(13): p. 3752-63 (2015).
24. Knight, A.K., et al., *An epigenetic clock for gestational age at birth based on blood methylation data*. Genome Biol. **17**(1): p. 206 (2016).
25. Bohlin, J., et al., *Prediction of gestational age based on genome-wide differentially methylated regions*. Genome Biol. **17**(1): p. 207 (2016).
26. Lee, H., et al., *DNA methylation shows genome-wide association of NFIX, RAPGEF2 and MSRB3 with gestational age at birth*. Int J Epidemiol. **41**(1): p. 188-99 (2012).
27. Engel, S.M., et al., *Neonatal genome-wide methylation patterns in relation to birth weight in the Norwegian Mother and Child Cohort*. Am J Epidemiol. **179**(7): p. 834-42 (2014).
28. Agha, G., et al., *Birth weight-for-gestational age is associated with DNA*

- methylation at birth and in childhood*. Clin Epigenetics. **8**: p. 118 (2016).
29. Ball, M.P., et al., *Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells*. Nat Biotechnol. **27**(4): p. 361-8 (2009).
 30. Cruickshank, M.N., et al., *Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy*. Genome Med. **5**(10): p. 96 (2013).
 31. Bibikova, M., et al., *High density DNA methylation array with single CpG site resolution*. Genomics. **98**(4): p. 288-95 (2011).
 32. Assenov, Y., et al., *Comprehensive analysis of DNA methylation data with RnBeads*. Nat Methods. **11**(11): p. 1138-1140 (2014).
 33. Aryee, M.J., et al., *Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays*. Bioinformatics. **30**(10): p. 1363-9 (2014).
 34. Morris, T.J., et al., *ChAMP: 450k Chip Analysis Methylation Pipeline*. Bioinformatics. **30**(3): p. 428-30 (2014).
 35. Chen, Y.A., et al., *Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray*. Epigenetics. **8**(2): p. 203-9 (2013).
 36. Fernando, F., et al., *The idiopathic preterm delivery methylation profile in umbilical cord blood DNA*. BMC Genomics. **16**: p. 736 (2015).
 37. Lin, X., et al., *Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome*. BMC Med. **15**(1): p. 50 (2017).
 38. Nordlund, J., et al., *Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia*. Genome Biol. **14**(9): p. r105 (2013).
 39. Triche, T.J., Jr., et al., *Low-level processing of Illumina Infinium DNA Methylation BeadArrays*. Nucleic Acids Res. **41**(7): p. e90 (2013).
 40. Teschendorff, A.E., et al., *A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data*. Bioinformatics. **29**(2): p. 189-96 (2013).
 41. Bakulski, K.M., et al., *DNA methylation of cord blood cell types: Applications for mixed cell birth studies*. Epigenetics. **11**(5): p. 354-62 (2016).
 42. Houseman, E.A., et al., *DNA methylation arrays as surrogate measures of cell*

- mixture distribution*. BMC Bioinformatics. **13**: p. 86 (2012).
43. Gelman, A., *Scaling regression inputs by dividing by two standard deviations*. Stat Med. **27**(15): p. 2865-73 (2008).
 44. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. Journal of the royal statistical society. Series B (Methodological): p. 289-300 (1995).
 45. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc. **4**(1): p. 44-57 (2009).
 46. Saxonov, S., P. Berg, and D.L. Brutlag, *A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters*. Proc Natl Acad Sci U S A. **103**(5): p. 1412-7 (2006).
 47. Farlik, M., et al., *DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation*. Cell Stem Cell. **19**(6): p. 808-822 (2016).
 48. Irizarry, R.A., et al., *The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores*. Nat Genet. **41**(2): p. 178-186 (2009).
 49. Yang, X., et al., *Gene body methylation can alter gene expression and is a therapeutic target in cancer*. Cancer Cell. **26**(4): p. 577-90 (2014).
 50. Nagarajan, R.P., et al., *Recurrent epimutations activate gene body promoters in primary glioblastoma*. Genome Res. **24**(5): p. 761-74 (2014).
 51. Mendizabal, I., et al., *Body-hypomethylated human genes harbor extensive intragenic transcriptional activity and are prone to cancer-associated dysregulation*. Nucleic Acids Res. **45**(8): p. 4390-4400 (2017).
 52. Grandi, F.C., et al., *Retrotransposition creates sloping shores: a graded influence of hypomethylated CpG islands on flanking CpG sites*. Genome Res. **25**(8): p. 1135-46 (2015).
 53. Bockmuhl, Y., et al., *Methylation at the CpG island shore region upregulates Nr3c1 promoter activity after early-life stress*. Epigenetics. **10**(3): p. 247-57 (2015).
 54. Joubert, B.R., et al., *Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns*. Nat Commun. **7**: p. 10577 (2016).

55. Joubert, B.R., et al., *DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis*. *Am J Hum Genet.* **98**(4): p. 680-96 (2016).
56. Zhang, H., et al., *Identification of methylation quantitative trait loci (mQTLs) influencing promoter DNA methylation of alcohol dependence risk genes*. *Hum Genet.* **133**(9): p. 1093-104 (2014).
57. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities*. *Mol Cell.* **38**(4): p. 576-89 (2010).
58. Vandenbroucke, L., et al., *Chorioamnionitis following preterm premature rupture of membranes and fetal heart rate variability*. *PLoS One.* **12**(9): p. e0184924 (2017).
59. Goldenberg, R.L., et al., *Epidemiology and causes of preterm birth*. *Lancet.* **371**(9606): p. 75-84 (2008).
60. Shen, M., et al., *Comparison of risk factors and outcomes of gestational hypertension and pre-eclampsia*. *PLoS One.* **12**(4): p. e0175914 (2017).
61. Moarii, M., et al., *Changes in correlation between promoter methylation and gene expression in cancer*. *BMC Genomics.* **16**: p. 873 (2015).
62. White, J.H., et al., *The GABAB receptor interacts directly with the related transcription factors CREB2 and ATFx*. *Proc Natl Acad Sci U S A.* **97**(25): p. 13967-72 (2000).
63. van Wassenaer-Leemhuis, A., et al., *Thyroid hormone supplementation in preterm infants born before 28 weeks gestational age and neurodevelopmental outcome at age 36 months*. *Thyroid.* **24**(7): p. 1162-9 (2014).
64. Starnawska, A., et al., *CACNA1C hypermethylation is associated with bipolar disorder*. *Transl Psychiatry.* **6**(6): p. e831 (2016).
65. Goonasekera, S.A., et al., *Decreased cardiac L-type Ca(2)(+) channel activity induces hypertrophy and heart failure in mice*. *J Clin Invest.* **122**(1): p. 280-90 (2012).
66. Harms, M.J., et al., *Prdm16 is required for the maintenance of brown adipocyte identity and function in adult mice*. *Cell Metab.* **19**(4): p. 593-604 (2014).
67. Yan, H., et al., *DNA methylation reactivates GAD1 expression in cancer by preventing CTCF-mediated polycomb repressive complex 2 recruitment*.

- Oncogene. **35**(30): p. 3995-4008 (2016).
68. Bonder, M.J., et al., *Disease variants alter transcription factor levels and methylation of their binding sites*. Nat Genet. **49**(1): p. 131-138 (2017).
 69. Benson, M.A., C.L. Tinsley, and A.J. Waite, *Ryanodine receptors are part of the myospryn complex in cardiac muscle*. Sci Rep. **7**(1): p. 6312 (2017).
 70. Nakagami, H., et al., *Gene polymorphism of myospryn (cardiomyopathy-associated 5) is associated with left ventricular wall thickness in patients with hypertension*. Hypertens Res. **30**(12): p. 1239-46 (2007).
 71. Xu, J., et al., *Investigation of Pathogenic Genes in Chinese sporadic Hypertrophic Cardiomyopathy Patients by Whole Exome Sequencing*. Sci Rep. **5**: p. 16609 (2015).
 72. Chen, X., et al., *GWA study data mining and independent replication identify cardiomyopathy-associated 5 (CMYA5) as a risk gene for schizophrenia*. Mol Psychiatry. **16**(11): p. 1117-29 (2011).
 73. Walczewska, J., et al., *The role of urocortins in the cardiovascular system*. J Physiol Pharmacol. **65**(6): p. 753-66 (2014).
 74. Diaz, I., et al., *miR-125a, miR-139 and miR-324 contribute to Urocortin protection against myocardial ischemia-reperfusion injury*. Sci Rep. **7**(1): p. 8898 (2017).
 75. Gagnon, K.B., et al., *Characterization of glial cell K-Cl cotransport*. Cell Physiol Biochem. **20**(1-4): p. 121-30 (2007).
 76. Hartmann, A.M., et al., *Differences in the large extracellular loop between the K(+)-Cl(-) cotransporters KCC2 and KCC4*. J Biol Chem. **285**(31): p. 23994-4002 (2010).
 77. Belenky, M.A., et al., *Cell-type specific distribution of chloride transporters in the rat suprachiasmatic nucleus*. Neuroscience. **165**(4): p. 1519-37 (2010).
 78. Chen, Y.F., et al., *Motor protein-dependent membrane trafficking of KCl cotransporter-4 is important for cancer cell invasion*. Cancer Res. **69**(22): p. 8585-93 (2009).
 79. Brown, T.C., et al., *DNA copy amplification and overexpression of SLC12A7 in adrenocortical carcinoma*. Surgery. **159**(1): p. 250-7 (2016).
 80. Mookherjee, N., et al., *Modulation of the TLR-mediated inflammatory response by the endogenous human host defense peptide LL-37*. J Immunol. **176**(4): p.

- 2455-64 (2006).
81. Chevrier, N., et al., *Systematic discovery of TLR signaling components delineates viral-sensing circuits*. *Cell*. **147**(4): p. 853-67 (2011).
 82. Xu, Y., et al., *The miR-184 binding-site rs8126 T>C polymorphism in TNFAIP2 is associated with risk of gastric cancer*. *PLoS One*. **8**(5): p. e64973 (2013).
 83. Rusiniak, M.E., et al., *Identification of B94 (TNFAIP2) as a potential retinoic acid target gene in acute promyelocytic leukemia*. *Cancer Res*. **60**(7): p. 1824-9 (2000).
 84. *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls*. *Nature*. **447**(7145): p. 661-78 (2007).
 85. Capsoni, S., et al., *Taking pain out of NGF: a painless NGF mutant, linked to hereditary sensory autonomic neuropathy type V, with full neurotrophic activity*. *PLoS One*. **6**(2): p. e17321 (2011).
 86. Breakefield, X.O., et al., *Structural gene for beta-nerve growth factor not defective in familial dysautonomia*. *Proc Natl Acad Sci U S A*. **81**(13): p. 4213-6 (1984).
 87. Naftelberg, S., et al., *Phosphatidylserine Ameliorates Neurodegenerative Symptoms and Enhances Axonal Transport in a Mouse Model of Familial Dysautonomia*. *PLoS Genet*. **12**(12): p. e1006486 (2016).
 88. Eyjolfssdottir, H., et al., *Targeted delivery of nerve growth factor to the cholinergic basal forebrain of Alzheimer's disease patients: application of a second-generation encapsulated cell biodelivery device*. *Alzheimers Res Ther*. **8**(1): p. 30 (2016).
 89. Tuszynski, M.H., et al., *Nerve Growth Factor Gene Therapy: Activation of Neuronal Responses in Alzheimer Disease*. *JAMA Neurol*. **72**(10): p. 1139-47 (2015).
 90. Dean, C., *What holds epigenetic memory?* *Nat Rev Mol Cell Biol*. **18**(3): p. 140 (2017).
 91. Iwasaki, M. and J. Paszkowski, *Epigenetic memory in plants*. *Embo j*. **33**(18): p. 1987-98 (2014).
 92. Lin, X., et al., *Choice of surrogate tissue influences neonatal EWAS findings*. *BMC Med*. **15**(1): p. 211 (2017).
 93. Huang, Y.T., et al., *Epigenome-wide profiling of DNA methylation in paired*

- samples of adipose tissue and blood. Epigenetics. 11(3): p. 227-36 (2016).*
94. Houde, A.A., et al., *Cross-tissue comparisons of leptin and adiponectin: DNA methylation profiles. Adipocyte. 3(2): p. 132-40 (2014).*