

ターゲットメチル化シーケンス

はじめに

ヒトゲノムの配列決定から数十年が経過し、研究者は遺伝的要因と多くの疾病状態の関連付けに取り組んできました。しかし、ゲノムDNAをコードするヌクレオチドは、遺伝子が細胞機能および全体的な健康に及ぼす影響の一部にしか過ぎません。エピジェネティクス（DNAの配列は変わらず起こる変化で、一部は遺伝する）も重要な役割を果たしています。DNAのメチル化は、シトシンとアデニンを化学的に修飾するエピジェネティックマーカーで、研究が進んでいます。シトシンのメチル化は、CpG部位と呼ばれるゲノムのCG配列で最もよく見られ、細胞特異的な方法で遺伝子発現を制御するために広く使用されています。^{1,2}ゲノムワイド関連解析(GWAS)はDNAのメチル化による変化を、がんや糖尿病など複雑な疾患および加齢などの複雑な生物学的状態や発生などと関連付けています。^{1,3-5}近年、メチル化シーケンスは遺伝子発現から広がり、疾患の診断分野で勢いが増しています。遺伝子発現に関わらず、メチル化の変動ががん検出にとって特に有益であり、感度が高いマーカーとなる研究例も示されています。⁶

メチル化シーケンスとキャプチャ

メチル化シーケンスでは、一連の脱アミノ化によってメチル化シトシンを変換せずに、メチル化されていないシトシンをウラシルに変換する酵素または化学的方法が用いられます（図1）。増幅時、ウラシルは相補鎖でアデニンと対になり、メチル化されていないシトシンの元の位置はチミンになります。最終産物は非対称であり、変換後は異なる2つの二本鎖DNA分子が生成されます（図1上段）。メチル化されているDNAに同じ処理をすることで、追加の配列セットが生じます（図1下段）。

キャプチャ前またはキャプチャ後の変換によって、ターゲットエンリッチメントを進めることができます。キャプチャ後の変換では、左側の元のサンプルDNAがターゲットとなります。キャプチャ前の変換では、右側に示した変換後配列の4本の鎖がターゲットとなります（図1）。キャプチャ後の変換では、プローブ設計の難易度は減りますが、PCR増幅でメチル化パターンが保持されず、キャプチャ前にはPCR増幅を実行できないため、たびたび多量のDNAインプット量が必要になります。したがって、セルフリーDNAなど、少ないインプットでの高感度アプリケーションでは、たびたびキャプチャ前の変換が適した手法となります。

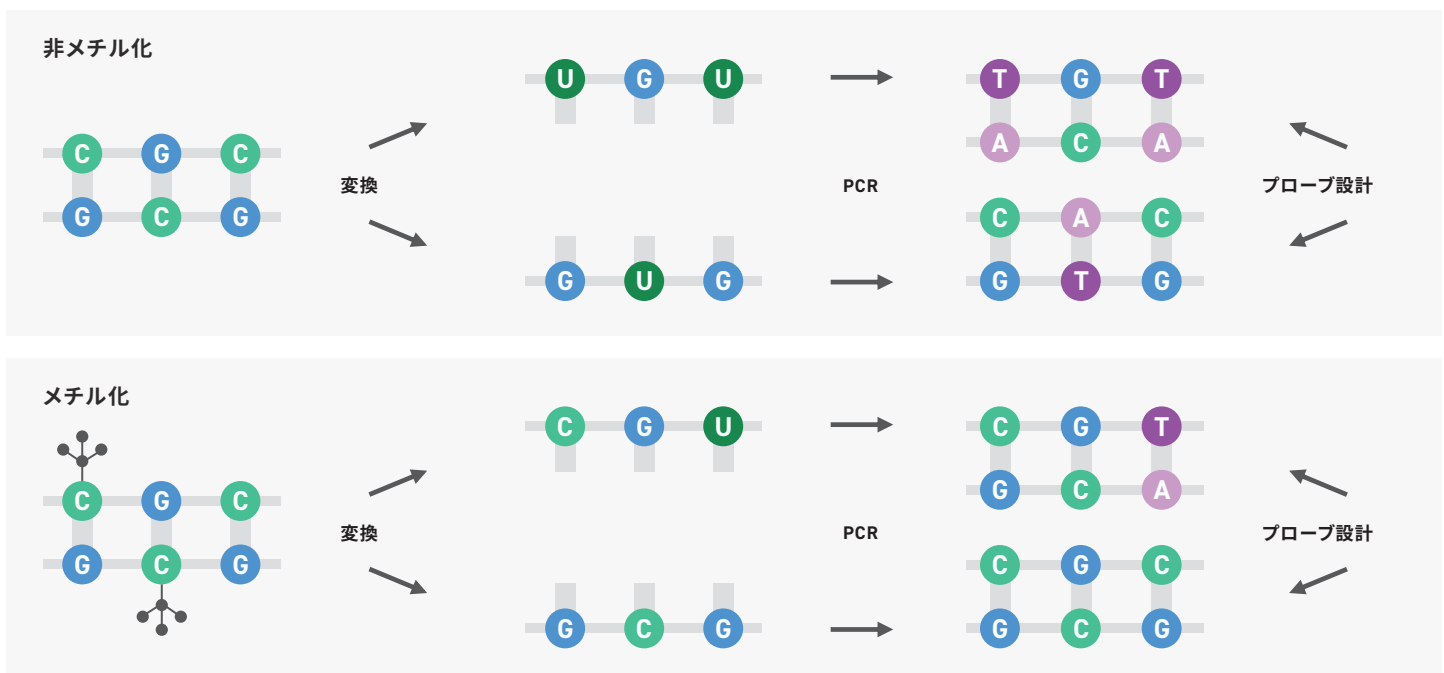


図1. メチル化変換。

パネル設計

キャプチャ前の変換では、変換後の4つの可能な配列のそれぞれをターゲットとするプローブが設計されるため複雑になります(図1)。実際には、ターゲットゲノムのコンテンツは数倍に増える一方で、メチル化されていないシトシンはチミジンに変換され、全体的なシーケンスの複雑度が低下します。この状況下の(配列空間)パネルで、しばしば高いオフターゲットとカバレッジ均一性の低下が起こるのはこのためです。

ここでは、Twistの高品質オリゴ合成、キャプチャソリューションおよび非常に均一なパネルが可能にするデータを根拠としたメソッドを活用して、ターゲットメチル化シーケンスの課題に取り組みました。Twistが新規に開発したメチル化特有のベースライン設計アルゴリズムと、ターゲットメチル化シーケンスのプロトコルを使用して、さまざまなサイズのパネルで高い性能の指標が実証されます。学習したデータとレスポンスを使用することで、さまざまレベルの厳密性で最適化された設計機能がデフォルトで実装されます。メチル化検出では、以前としてアレイが広く使用されていますが、高品質のターゲットエンリッチメントパネルは、細胞固有のダイナミックなメチル化ターゲットの探索、あるいはまだ十分に解明されていない非翻訳領域をターゲットとした探索を一塩基対の解像度で行うことで、静的なアレイの設計に代わる魅力的な手段となります。

材料と方法

NA12878 (Coriell Institute)のゲノムDNAサンプルおよびEpiScope®低メチル化と高メチル化gDNAコントロール(それぞれ5%未満および95%超メチル化HCT116 DKO gDNA)を(Covaris® ME220で)機械的に約300 bpのサイズに断片化しました。断片化された低メチル化および高メチル化コントロールを混合して、さまざまにシミュレートされたメチル化レベルのサンプルを調製しました。Swift Accel-NGS® Methyl-seq DNAライブラリキットに500 ngのインプットgDNAを適用し、バイサルファイト処理(Zymo EZ DNA Methylation-Lightningキット、Omega Bio-Tek Mag-Bind RxnPure Plus SPRI Beads、KAPA HiFi Uracil+ DNA Polymerase)と組み合わせました。NEBNext® Enzymatic Methyl-seqキットには、200 ngのインプットgDNAを適用しました。断片化されたサンプルとライブラリは、Agilent BioAnalyzer 7500とInvitrogen Qubit Broad Rangeキットを用いて検証しました。

Twistの4時間高速ハイブリダイゼーションプロトコルは、Twistメチル化パネルで使用するために見直しを行いました。さまざまなターゲットサイズ(0.05、1.0、1.5、3.0 Mb)の範囲をカバーする4つのメチル化パネルを合成し、性能評価に用いました。各シングルプレックスキャプチャに200 ngのライブラリを使用し、その後Illumina® NextSeq® 550でv2.5 High Outputキットを使用して2x151 bpでシーケンスしました。サンプルあたり250xのRawカバレッジにサンプリング後、アライメントおよびメチル化解析をPicardHsMetricsとBismark 19.1を用いて実施しました。

結果

さまざまなサイズのパネルによるテスト

キャプチャ前の変換は、非常に高感度なエピジェネティックアプリケーションを可能にしますが、主な課題は変換後のゲノムの多様性の低下です。非メチル化パネルと比較して、一般的に著しく高いオフターゲット(>50-60%のレベル)、ターゲットカバレッジ低下および、キャプチャ均一性の大きな低下(fold 80ベースペナルティ値は>2.5)を導きます。Twistのターゲットメチル化シーケンスプロトコルは、さまざまなメチル化ターゲットを広範囲にカバーする異なる4つのサイズのパネルを使用してテストされました。Twistのベースラインパネル設計機能を使用すると、その後の実験で使用したパネルは、オフターゲット値27%という現在の業界標準を下回るオフターゲット水準を示しました。0.05 Mbパネルは、他の3つのパネルと比較して、高いオフターゲットを示しました。この結果はターゲットサイズが非常に小さい性質による可能性があります。とりわけこのアプリケーションではキャプチャ均一性も、業界標準である>2.5のFold-80ベースペナルティより優れており、1.75および1.5という値に達しました。テストした4つすべてのパネルで、duplicate率は非常に低く、キャプチャ過程が非常に効率的で、ワークフローを通して高いサンプル複雑性を維持できることを示しました。全体として、250xのRawシーケンスにより、最小のパネルでも30xで70%を超えるカバレッジを達成しました。

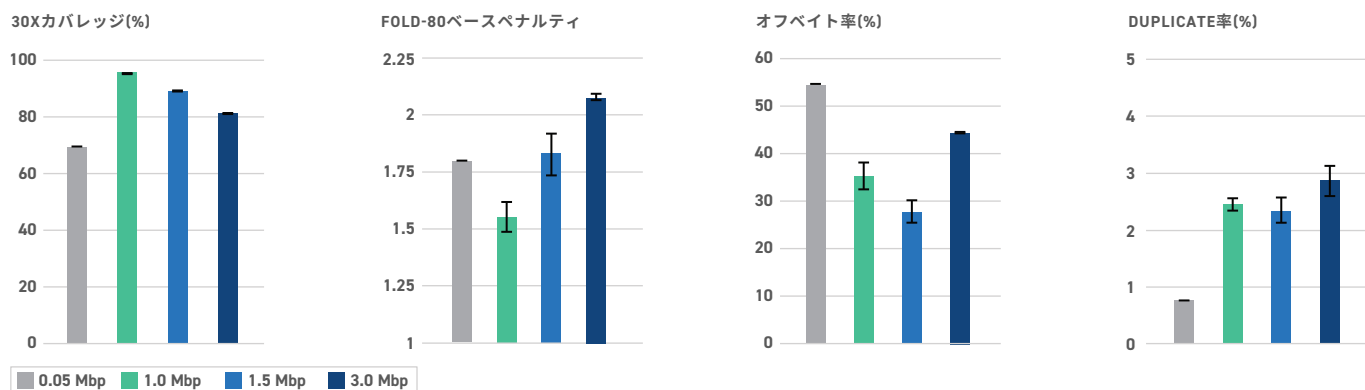


図2. パネルサイズ。

高品質な事前最適化された設計による厳密な制御

Twistのパネル設計適応最適化アルゴリズムにより、キャプチャ実験の経験的なデータを使用してプローブ特異性について学習し、パフォーマンスを定量的にチューニングできます。この強力な手法は、オフターゲット率の制御を優先するメチル化パネルで特に有効です。さらに、3万以上のメチル化ターゲットで収集したデータを使用して、有益なシーケンスの特徴を取得し、これらを使用してカスタムパネルに3段階の厳密性を持たせてデフォルトのパネル設計が出来るように最適化を行い開発しました。デフォルトパネル設計の実証用に、低、中、高の厳密さで設計された1 Mbパネルを使用しました。その結果、オフターゲット率の制御が強化される一方で、他の主要な指標はほとんど変化しませんでした（図3）。

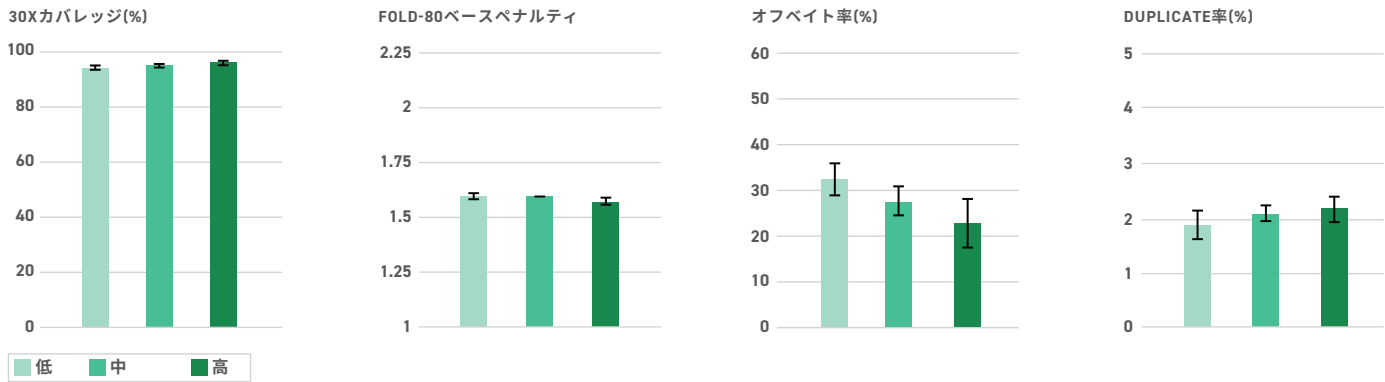


図3. 設計厳密性。

キャプチャ性能はターゲットの想定された低メチル化から高メチル化まで維持

想定されたメチル化レベルでの適合性を評価するため、低メチル化と高メチル化から生成された細胞株のgDNAライブラリを各0、25、50、75、100%メチル化の最終比率となるように混合し、「中」の厳密性で1 Mbパネルでキャプチャを実行しました。以下の図で、主要なキャプチャ指標をグラフに示しました。グラフの棒は、メチル化の程度が異なるサンプル間のキャプチャ性能の変動を示しています。これらの指標は、メチル化レベルの変化に対して無反応～弱い反応であり、低メチル化および高メチル化DNAを含む幅広いメチル化状態に応じたこのシステムの適合性が確認されました。

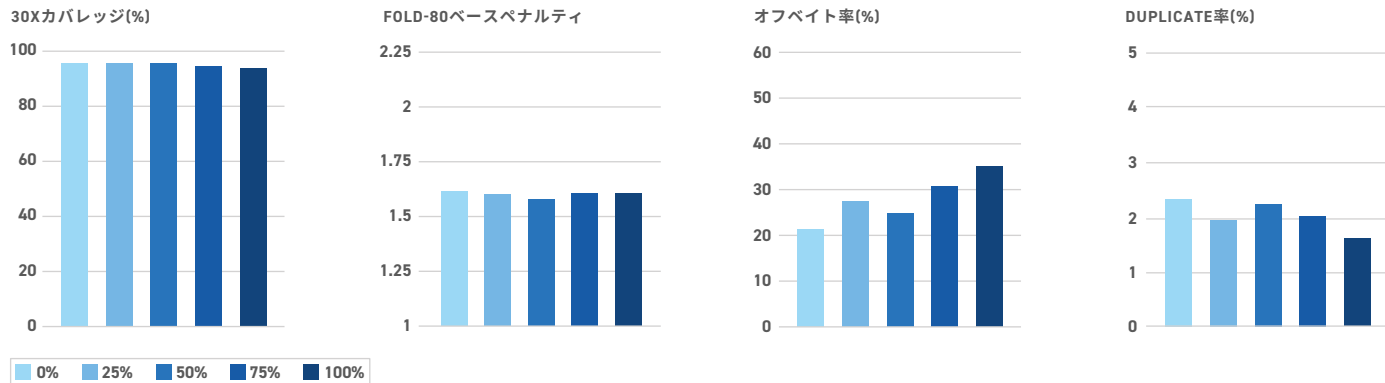


図4. %メチル化ライブラリー。

メチル化レベルの検出

プロモーターやその他の調節因子のメチル化レベルの変化を、がんの早期発見のための高感度のマーカーとして利用できないか、研究が進んでいます。⁷ターゲットメチル化シーケンスがさまざまなレベルのDNAメチル化を検出および定量化できることを示しました。低メチル化および高メチル化DNAをさまざまな割合で混合し、1 Mbパネルを用いてキャプチャに使用しました。図5は、乳がんなど特定のがんでメチル化状態が変化することが知られている臨床的に関連のあるCyclin D2遺伝子座のターゲットおよび個別のCpG部位に沿った異なるDNAメチル化レベルの検出を示しています。⁸



図5. メチル化検出の感度。

酵素およびバイサルファイトベースアッセイとの適合性

メチル化シトシンの検出では、非メチル化シトシンのチミンへの変換が行われますが、メチル化シトシンは変換されません。従来の変換はバイサルファイトを使った化学的な手法で行われていました。この分野では非メチル化シトシンの酵素による変換など、(これまでに提示された結果で使用されている) その他の手法が使われることが増えています。変換反応条件に対する酵素の感受性が高いことや、バイサルファイトによるDNAの偏った分解など、変換の方法ごとに利点や欠点があります。

Twistのターゲットメチル化シーケンスは、酵素およびバイサルファイトの両方の手法に適合しています(図6)。非CpG部位で変換されたシトシンの割合として測定された変換率は両方の方法でも99.5%超でした(data not shown)。総合的なキャプチャ指標は両方のライブラリ調製法でほぼ同じ程度でしたが、バイサルファイト法では均一性とオフターゲットなど特定の指標が低下しました。均一性の低下は、少なくとも部分的にはバイサルファイト法によるライブラリ調製で導入された固有のGCバイアスによって起こったことを示す証拠が得られています(data not shown)。

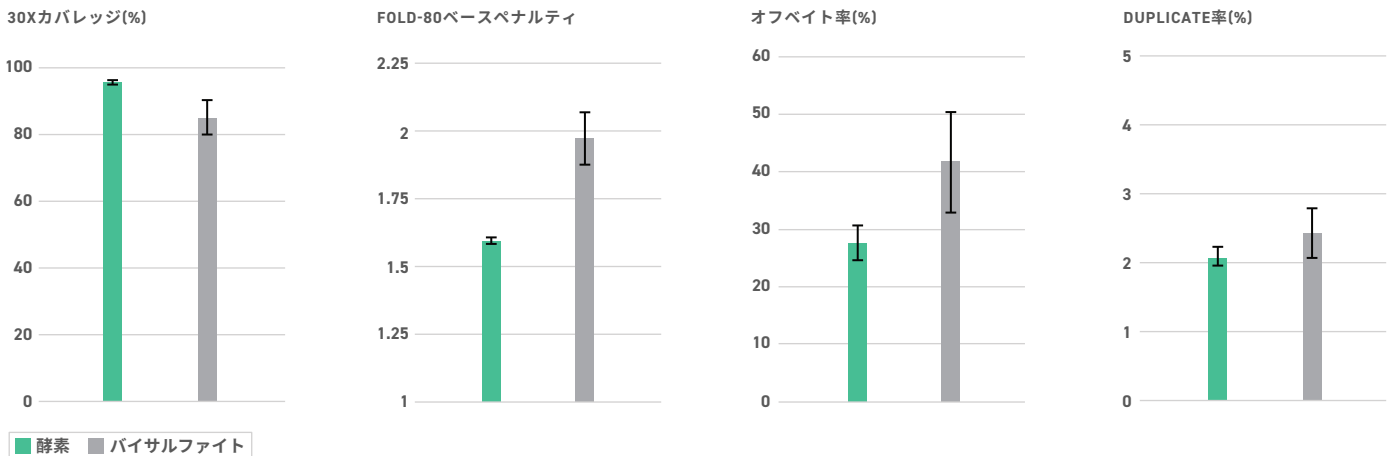


図6. ライブラリ調製におけるメチル化変換法。

結論

結果として、メチル化シーケンスのターゲットエンリッチメントパネルにおけるTwistテクノロジーのユニークな特性が実証されました。広範囲のメチル化ターゲットとレベルでの優れた性能と、さまざまなレベルの設計厳密性から選択できる柔軟性を備えた、高度に最適化された従来の常識にとられない設計が可能です。バイサルファイトにも酵素変換法にも適合するTwistの高い均一性メチル化パネルは、高感度エピジェネティックアプリケーションの開発の限界に挑む新しいツールです。

参考文献

1. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13, 484–492 (2012).
2. Yang, X., Shao, X., Gao, L. & Zhang, S. Comparative DNA methylation analysis to decipher common and cell type-specific patterns among multiple cell types. *Brief Funct Genomics* 15, 399–407 (2016).
3. Loh, M., Zhou, L., Ng, H. K. & Chambers, J. C. Epigenetic disturbances in obesity and diabetes: Epidemiological and functional insights. *Mol Metab* 27S, S33–S41 (2019).
4. Ferrucci, L. et al. Measuring biological aging in humans: A quest. *Aging Cell* 19, e13080 (2020).
5. Pérez, R. F. et al. Longitudinal genome-wide DNA methylation analysis uncovers persistent early-life DNA methylation changes. *J Transl Med* 17, 15 (2019).
6. Locke, W. J. et al. DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Front Genet* 10, (2019).
7. Hao, X. et al. DNA methylation markers for diagnosis and prognosis of common cancers. *PNAS* 114, 7414–7419 (2017).
8. Evron, E. et al. Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation. *Cancer Res.* 61, 2782–2787 (2001).