

パネル設計およびエンハンサー試薬による ターゲットメチル化シーケンスにおける オフターゲットキャプチャの最小化

はじめに

メチル化シーケンスおよびキャプチャ

メチル化シーケンスでは、酵素または化学的方法を用いて、非メチル化シトシンをウラシルに変換します。いずれの方法でも、メチル化シトシンは変換されません(図1)。増幅時に各ウラシルは相補鎖上のアデニンと対になり、その結果、非メチル化シトシンの元の位置にチミンが組み込まれます。最終産物は非対称性であり、変換後に異なる2つの二本鎖DNA(dsDNA)分子が生成します(図1上段)。メチル化DNAに対して同じ処理をすることにより、さらに2つのdsDNA分子が生成します(図1下段)。

ターゲットエンリッチメントは、変換プロセスの前か後で実施することができます。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ないため、たびたび多量のインプットDNA材料が必要になります。したがって、微量のセルフリーDNAを測定するリキッドバイオシーなど、少ないインプットでの高感度アプリケーションでは、多くの場合、キャプチャ前の変換(変換後に増幅可能)が最適な手法となります。しかしキャプチャ前の変換では、最適なキャプチャのために変換後に予想される配列4つそれぞれをターゲットとするプローブを設計しなければならないため、より複雑なパネルが必要となるなど、いくつかの課題があります(図1)。

シーケンスの複雑性に対する変換の影響

従来、メチル化シーケンスライブラリのハイブリッドキャプチャでは、オフターゲットキャプチャの増加とキャプチャ均一性の低下が生じます。これらの問題が起こるのは、非メチル化シトシンからチミンへの変換により、ターゲットゲノムにおける配列の複雑度やGC含量が低下するためです(図2)。

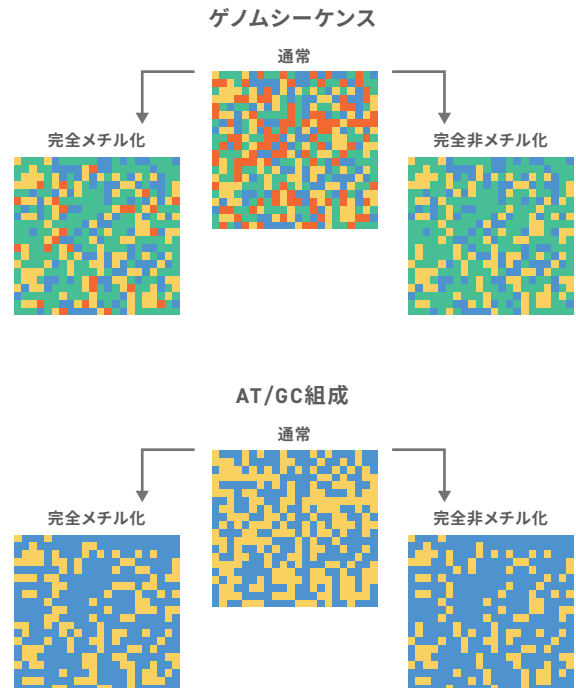


図2: シーケンスの複雑度およびGC含量に対する変換の影響。ゲノムの通常のヌクレオチド組成と比較して、メチル化シトシンの変換では、シーケンスの複雑度が著しく低下します。上記の例はランダムなゲノムシーケンスを示しています。4つの色は4つの各ヌクレオチドを表し(上)、2つの色はG/CまたはA/Tを表しています(下)。変換後のシーケンスの複雑度とGC含量(%)の低下はGCリッチな機能的配列やメチル化ターゲットにおいて顕著であり、繰り返し配列の増加、オフターゲットキャプチャ、並びに融解温度やシステムの化学反応速度の広範な変化などの課題を生じます。

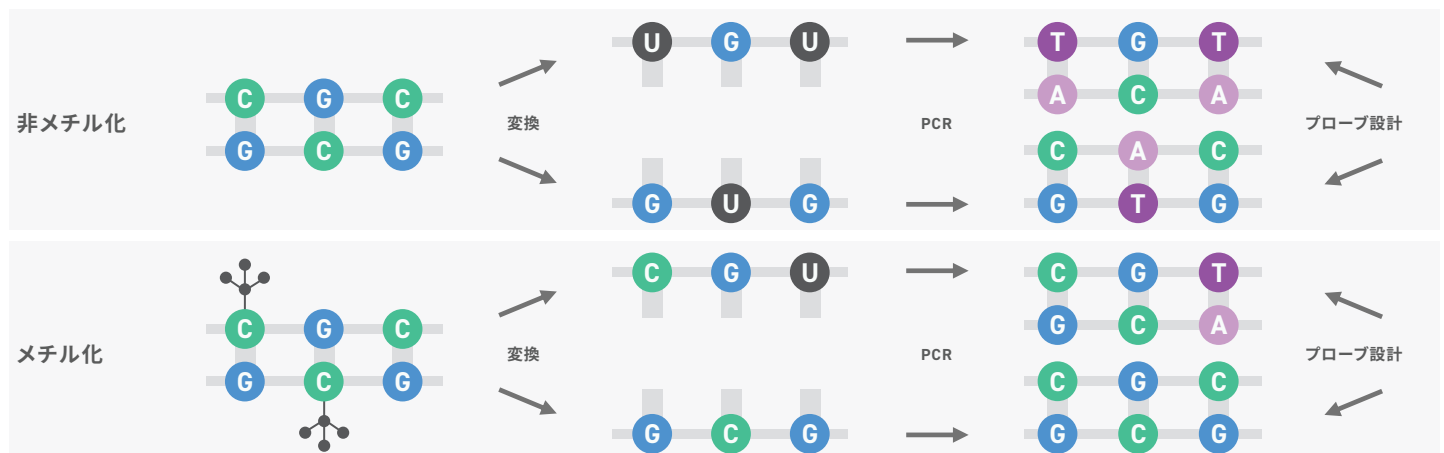


図1: メチル化変換のプロセス。非メチル化DNAとメチル化DNAの変換時の配列の違い。

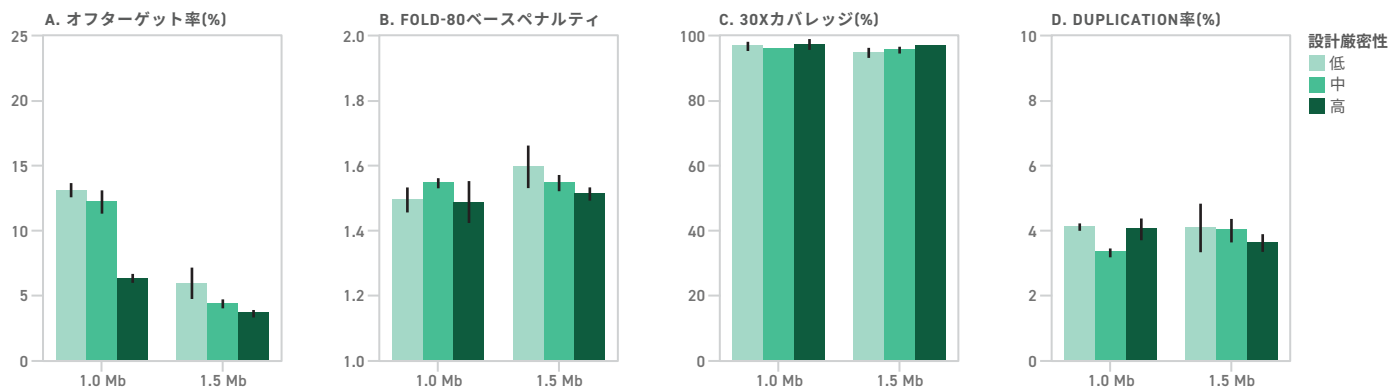


図3：シーケンス性能に対する設計の厳密性の影響。設計の厳密性が増すと、オフターゲットキャプチャが減少します。1Mbと1.5Mbのカスタムメチル化パネルおよび Twist Targeted Methylation Sequencing Protocol を用いて、厳密性が低、中、高の設計を比較しました。以下の推奨キャプチャ条件を使用：メチル化エンハンサー 2 uL、洗浄バッファー 1 の温度 65°C、ハイブリダイゼーション時間 2 時間。このテクニカルノートで得られた結果は、1) NextSeq® 500/550 高出力 v2 キットを用いたシーケンスによる、2x76 のペアエンドリードの生成、2) パネルの標的サイズに対して 200x に調整したカバレッジへのダウンサンプリング、Bismark Aligner を用いたマッピング、および Picard HS Metrics の使用によるものです。

Twist Bioscience は、ターゲットメチル化シーケンスの課題を克服するべく、前例のないレベルのオフターゲット制御とキャプチャ性能を可能とするメチル化パネルを用いた以下の 2 つの相補的な革新的手法を取り入れました。

- 異なるレベルのフィルタリング厳密性で最適化されたパネル設計。性能を微調整するためにプローブを取り除いたり、困難な領域の重要なターゲットをカバーできるようにプローブを維持したりなど、厳密性のレベルを制御。
- Twist メチル化エンハンサー（変換後のキャプチャシステムの効率を最大限に高める新しい試薬）。

上記の手法を開発するために、数万のメチル化ターゲットからデータを収集し、変換されたゲノムの複雑度、配列の繰り返しおよび配列の前後関係がどのようにパネル性能に影響するか検討しました。フィルタリングに使用したいいくつかの従来のメトリクスは全て適用しました。このような根拠から、有益なシーケンスの特徴を導き出し、これを利用して、このテクニカルノートに示したパフォーマンスの向上を支える新たな最適化手法を開発しました。

メチル化検出では依然として静的なマイクロアレイが広く使用されていますが、このテクニカルノートで説明する革新的手法により可能となった高品質のターゲットエンリッチメントパネルは、一塩基解像度で動的かつ細胞特異的なメチル化ターゲットを探索するための魅力的な代替手段となります。

結果

パネル設計によりオフターゲットキャプチャを低減

変換後のゲノムの複雑度は変化するため、メチル化パネル用のプローブ配列のフィルタリングは、基準の変換されていないゲノムよりもかなり困難なものとなります。メチル化検出ライブラリでは、低複雑度、低 GC、繰り返し配列のため、ユーザーは、ターゲット領域全体でオフターゲットキャプチャの増加、カバレッジの低下、バリエーションの信頼性低下の可能性がある中で、検討するターゲットの含有バランスを取らなければなりません。

つまり、多くの場合、パネル設計の改善のためには経験的なキャプチャデータを用いて繰り返し実施する必要があり、ユーザーに負担がかかります。この解決策として、カスタムメチル化パネルの設計時に

フィルタリングの厳密度を異なる 3 つのレベルに設定して、従来の常識にとらわれない高度に最適化されたターゲットエンリッチメントを実現する手法を開発しました。これにより、ユーザーは、発注時に要件に応じてターゲットとフィルタリングの厳密性を調整することができます。

デフォルトパネル設計のテスト用に、低、中、高のフィルタリング厳密性で設計された 1 Mb と 1.5 Mb の 2 つの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ことが示されました。中度の厳密性ではプローブを 1% 未満除去し、高度ではプローブを 3 ~ 5% 除去した場合、オフターゲットは 4% もの低さとなりました。重要な点は、その他の指標は設計の厳密性が変化しても概ね影響を受けなかったことです (図 3)。

メチル化エンハンサーによるオフターゲットキャプチャの低減

ターゲットエンリッチメントシステムには、通常、キャプチャ時のクロスハイブリダイゼーションを最小限に抑えるブロッカーが含まれています。ハイブリダイゼーション反応にブロッカーを加えることで、非特異的な DNA 相互作用が制限されますが、それでも最終的に標的としにくいライブラリ配列のプルダウンが起こります。市販のブロッカーは標準的なターゲットエンリッチメントのアプリケーションにおけるオフターゲットキャプチャの低減には効果的ですが、メチル化シーケンス DNA ライブラリのハイブリッドキャプチャには最適化されていません。Twist Bioscience は、ターゲットメチル化シーケンス用に特別に設計した、メチル化エンハンサーと呼ぶ独自のブロッカーセットを作成しました。本製品により、酵素やバイサルファイト変換によって生成される繰り返し配列からなるオフターゲットライブラリのキャプチャが効果的に抑制されます。なお本製品は、他の重要なハイブリッドセレクションの指標に悪影響をもたらすことはありません。

Twist のメチル化エンハンサーは、ハイブリッドセレクションの指標に悪影響を及ぼすことなく、カスタムパネルのターゲット領域およびインプットゲノム DNA のメチル化状態に応じて、様々な方法でオフターゲットキャプチャを低減させます。シングルプレックス反応において、厳密性が中度の 1.0 Mb と 1.5 Mb のメチル化パネルを用いて、Twist のターゲットメチル化シーケンスワークフローに Twist メチル化エンハンサーを適用しました。その結果、オフターゲット率は 2 倍も低下し、オフターゲット値は 4% もの低値でした。均一性 (fold 80

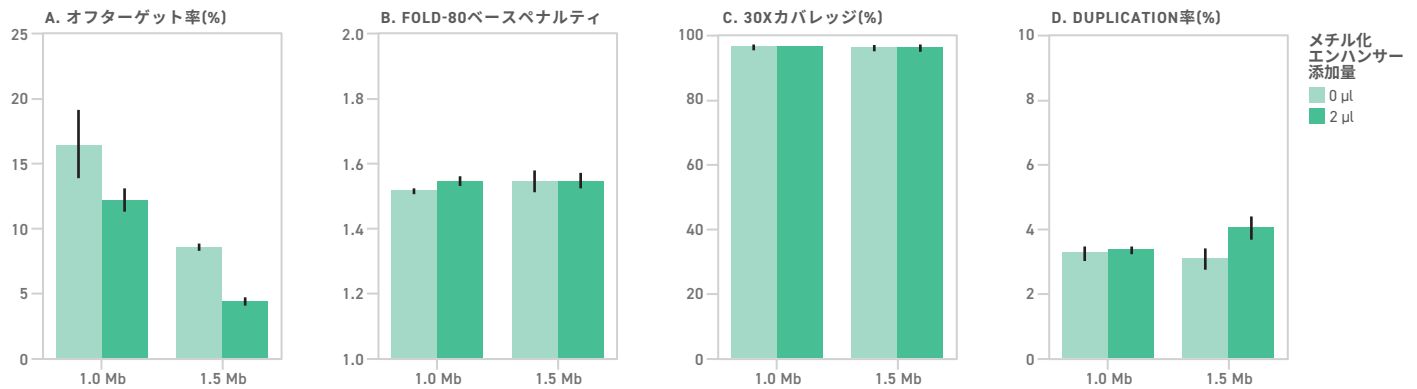


図 4: シーケンス性能に対する Twist メチル化エンハンサーの効果。Twist メチル化エンハンサーは、非特異的なオフターゲットキャプチャをブロックすることで、オフターゲットキャプチャを効果的に抑制します。厳密性中程度で設計した 1.0 Mb と 1.5 Mb のカスタムメチル化パネルおよび 200 ng のライブラリ (NA12878, Coriel) を用いたハイブリッドキャプチャを実施しました。Twist が推奨するキャプチャ条件 (Twist メチル化エンハンサー 2 μ L、洗浄バッファー 1 の温度 65°C、ハイブリダイゼーション時間 2 時間) をすべての反応で使用しました (その他の方法については図 3 を参照)。

ベースペナルティ)、30x カバレッジ、Duplicate 率などのその他の主要なキャプチャ指標には、概ね変化はありませんでした (図 4)。

従来型のブロkkerはメチル化に対して最適に調整されていませんが、変換後の配列に対しては依然としてブロッキング活性を示しています。以下のように、メチル化 DNA に対する Twist のメチル化エンハンサーのブロッキング性能を標準的なブロkkerと比較しました。最適化されていない 1.5 Mb のカスタムメチル化パネルを用いてキャプチャしたシングルプレックスのハイブリダイゼーション反応において、両方のブロkkerを個別に量を変え、その性能を比較しました。比較のために各々同じ量のブロkkerを追加しました。この場合、Twist メチル化エンハンサーではオフターゲットキャプチャは約 40% 低下しましたが、従来型の標準的なブロkkerでは約 15% の低下でした (図 5)。Twist Bioscience は、各ハイブリダイゼーション反応においてメチル化エンハンサーを 2 μ L から開始し、各カスタムメチル化パネルの必要性に応じて最大 5 μ L まで増量するよう推奨しています。

結論

非メチル化シトシンの変換によりゲノムシーケンスの複雑度が低下するため、インプット量の少ない検体におけるターゲットメチル化シーケンスは以前から大きな課題となっていました。Twist のメチル化検出システムは、異なるレベルの最適化されたフィルタリング厳密性を実現するパネル設計機能と、オフターゲットの低減によりシステム性能を改善するよう相乗効果を発揮するように仕立てたブロkkerであるメチル化エンハンサーの導入により、こうした障害を効果的に克服します。これらの両方の最適化により、メチル化検出用パネルで認められたシーケンス性能を大幅に改善し、リキッドバイオシーなどの高感度アプリケーションを可能とする Twist のメチル化検出システムの基礎となる革新的手法の一つとなります。

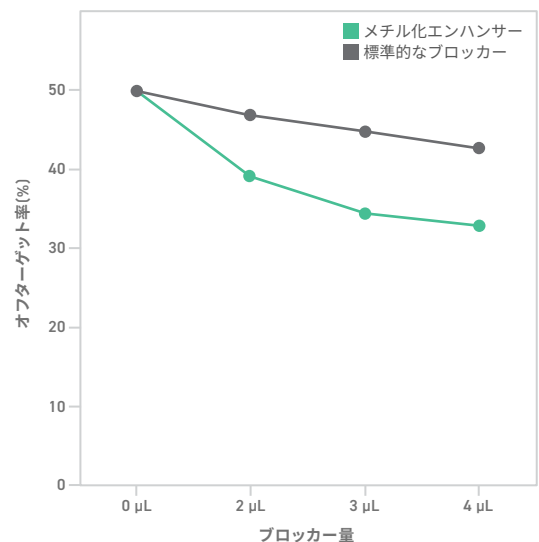


図 5: Twist メチル化エンハンサーおよび標準ブロkkerの量の影響。メチル化エンハンサーは、同じ量の添加で標準的なブロkkerよりもオフターゲットをより効果的に低減します。上記のデータは、2 つのブロkkerシステムを用いて別々に生成しました。各反応では、各ブロkkerを同じ量使用しました。ハイブリッドキャプチャは、1.5 Mb の最適化されていないカスタムメチル化パネルおよび 200 ng のライブラリ (NA12878, Coriel) を使用して実施しました。その他の方法については図 3 を参照。