

イルミナCMOS Chipおよび1色法SBSケミストリー

iSeq™ 100システムはCMOSテクノロジーと革新的な1色法SBSケミストリーを組み合わせて、コンパクトなシステムで高精度なデータを生成

はじめに

イルミナシーケンスプラットフォームは世界中で最も広く採用されている次世代シーケンスケミストリーであり、実証された高精度な1塩基合成（SBS）ケミストリーを採用しています。¹ イルミナSBSは、独自の可逆的ターミネーター法を使用して、伸長するDNA鎖に取り込まれる1塩基を検出し、大量並列シーケンスを可能にします。イルミナSBSケミストリーは4つ全てのヌクレオチドの自然競争により、取り込みのバイアスを減らし、繰り返し領域やホモポリマーのさらに安定したシーケンスを実現します。² キャピラリー電気泳動によるSangerシーケンスと比較して、NGSはより迅速により少ないハンスオンステップで、低頻度バリエーションや隣接したフェージングバリエーション（相同染色体別のバリエーション）など、幅広いDNAバリエーションを検出することができます。^{3,4}

SBSケミストリーは全てのイルミナシーケンスプラットフォームの基盤となっていますが、光学システムの進化により、4色、2色、1色で検出する方法が開発されています。iSeq 100システムは、相補型金属酸化膜半導体（CMOS）チップ上で実施する1色法SBSを用いたシーケンスを行うイルミナSBSテクノロジーの最新進化システムです。1色法SBS法を用いることで、低いシーケンスコストと、たった1立方フィート（0.03立方メートル）の装置サイズという便利さを実現しています。

高精度なSBSケミストリー

イルミナSBSケミストリーは可逆的に標識されるヌクレオチドを用いて、1つのフローセル表面で何億ものクラスターを並行にシーケンスします（図1）。各シーケンスサイクル中に、標識デオキシヌクレオチド三リン酸（dNTP）1個が核酸鎖に加えられます。標識ヌクレオチドは重合化のターミネーターとして働き、各dNTP取り込みの後、蛍光色素が画像化され、その後次のヌクレオチドの取り込みができるように化学的に処理されます。各サイクル中のシグナル強度の測定からベースコールが行われます。可逆的ターミネーターに結合した4種類のdNTP（A、C、T、G）が単一の分子として存在するため、自然競争により取り込みバイアスが最小に抑えられます。さらに、可逆的ターミネーションによって、シーケンスサイクルごとに鑄型鎖当たり1塩基しか取り込まれないため、ホモポリマー領域でのシーケンス精度が向上し、その他のテクノロジーと比較してエラー率が低下します。²



イルミナSBSテクノロジーのアニメーションビデオはこちらをご覧ください。
jp.illumina.com/SBSvideo

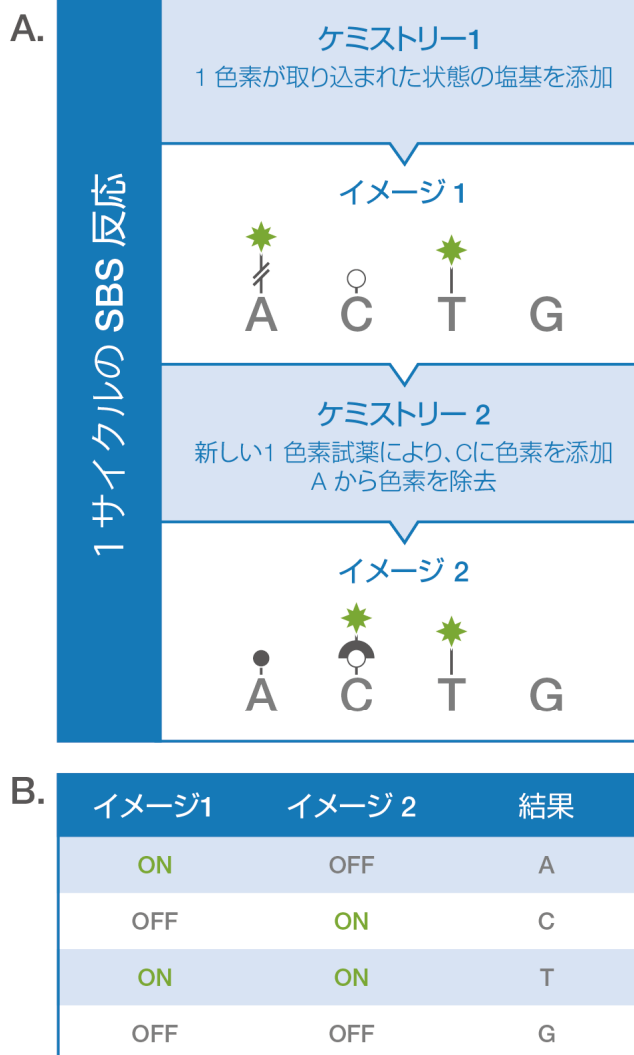


図 1: 1色法検出によるイルミナSBSケミストリー (A) iSeq 100システムは革新的な1色法ケミストリーを採用しています。各シーケンスサイクルには2回のケミストリーステップと2回のイメージングステップが含まれます。初めのケミストリーステップでは、蛍光標識されたアデニンとチミンを含むヌクレオチド混合物がフローセルにロードされます。初めのイメージングステップ中に、各クラスターからの励起光がCMOSセンサーによって記録されます。2回目のケミストリーステップでは、アデニンから蛍光標識を取り除き、チミンに蛍光標識を付加します。両方のケミストリーステップでグアニンは暗いままです（非標識）。その後2回目のイメージが記録されます。(B) イメージ1とイメージ2の組み合わせがイメージ解析ソフトウェアで処理され、どの塩基が各クラスターポジションに取り込まれたかを同定します。このシーケンスサイクルはn回繰り返され、n塩基のリード長を解析します。

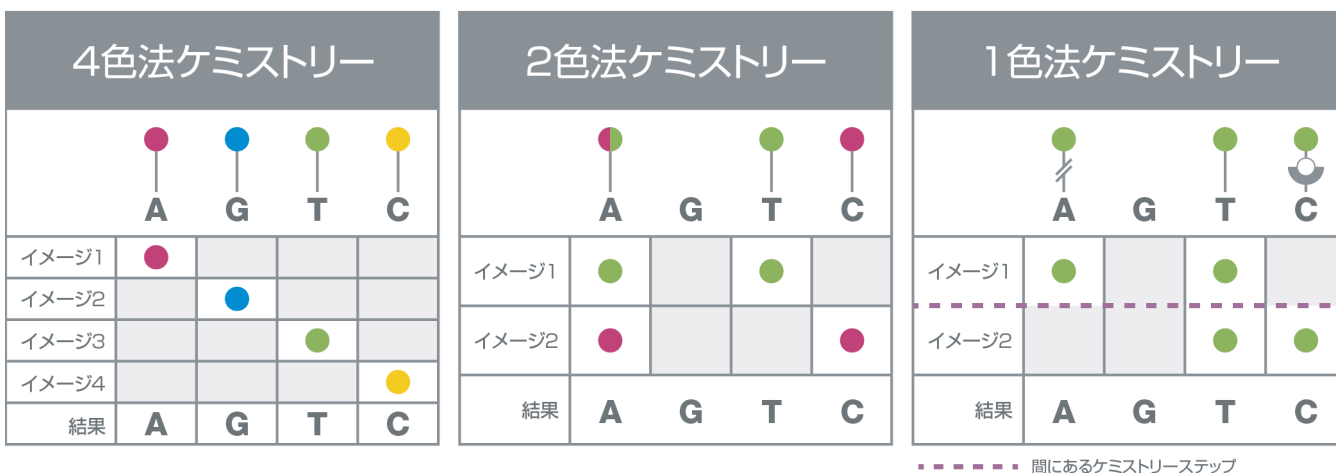


図 2: 4色、2色、1色法ケミストリー — 4色法ケミストリーは4色の異なる蛍光色素による標識ヌクレオチドの混合物を使います。2色法ケミストリーは2色の異なる蛍光色素を使い、1色法ケミストリーは1色の色素だけを使います。イメージはイメージ解析ソフトウェアで処理され、ヌクレオチドの同定が行われます。

4色法SBSケミストリー

4色法SBSでは、塩基は各塩基に対して4色の異なる蛍光色素を使い、シーケンスサイクルごとに4イメージを取得して同定されます (図2A)。4色の異なる標識の全塩基をフローセルに加えるケミストリーステップにより、シーケンスサイクルが開始されます。ヌクレオチドの取り込みに続いて、イメージングサイクルが開始され、このサイクルでは4つの異なる光波長を使って4色別々のイメージを取得します。イメージはイメージ解析ソフトウェアで処理され、どのヌクレオチドがフローセル中の各クラスターポジションで取り込まれたかを特定します。そのため、4色法シーケンスでは、各シーケンスサイクルはDNA配列を特定するために4色の色素と4イメージが必要になります。MiSeq™およびHiSeq™シリーズシステムは4色法SBSを採用しています。

2色法SBSケミストリー

各塩基に別々の色素を使うのではなく、2色法SBSは、4塩基のベースコールを特定するために、2色の蛍光色素と2つのイメージでヌクレオチドの検出を単純化しています (図2B)。イメージの取得は赤と緑のフィルターを用いて行われます。チミジンは緑の蛍光色素で標識、シトシンは赤の蛍光色素で標識、アデニンは赤と緑の蛍光色素で標識します。グアニンは常に暗いままです。MiniSeq™、NextSeq™、およびNovaSeq™システムは2色法ケミストリーを採用しています。

1色法SBSケミストリー

iSeq 100 システムは実証済みのイリミナSBSケミストリーとCMOSテクノロジーを組み合わせ、1色法シーケンスケミストリーを実現したシステムです。このシステムはCMOSチップ上に作製されたナノウェルによるパターン化フローセルを採用しています (図3)。光ダイオード (pixel) 上に直接並べられたナノウェル内で、クラスタリングとシーケンスが行われます。消耗品に組み込まれたCMOSセンサーを使うことで、シンプルかつ迅速な検出が実現しています。

各ヌクレオチドに異なる色素を使う4色法SBSケミストリーと異なり、iSeq 100 システムは1色の色素を使い、シーケンスサイクルごとに2回のケミストリーステップと2回のイメージングステップを行います。1色法SBSケミストリーでは、アデニンの標識は取り除くことができ、初めのイメージの時だけ標識されています。シトシンは蛍光標識と結合できるリンカーグループを保有しており、2回目のイメージの時だけ蛍光が検出されます。チミジンは常に蛍光標識が検出されるため両方のイメージングで蛍光標識が検出され、グアニンは常に暗いままです。ヌクレオチドは2つのイメージ中の各塩基に対する異なる励起パターンを解析することで同定されます (図2C)。

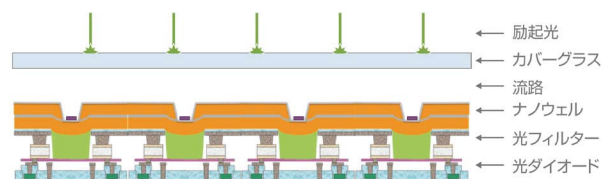


図 3: イリミナCMOSフローセルの構造 — シーケンスライブラリーはiSeq 100試薬カートリッジにロードされ、このカートリッジにはCMOSチップ上に作製されたパターン化フローセルが含まれています。フローセルの各ウェルはCMOS光ダイオード上に並べられています。クラスター形成中に、専用のExAmpケミストリーによってフローセル中の各ウェルは単一のクローンクラスター形成を行います。各イメージングステップの間、励起光はCMOS光ダイオードによって検出されます。

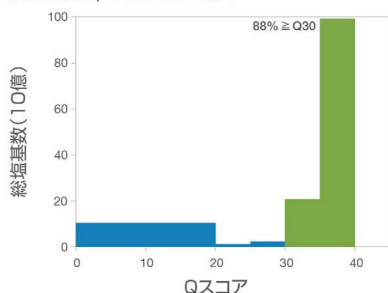
表 1: バリエント解析の装置間比較

システム	オンターゲット率 (%) (アラインされたリード)	ミスマッチ率 (%)	カバレッジの 均一性率 (%) ^a	SNPの正確性 (%)	SNPのリコール率 (%)
パネル1 (BRCA) ^b					
iSeq 100	96.0	0.23	100	90.5	100
MiniSeq	95.8	0.21	100	89.1	100
MiSeq	95.8	0.20	100	89.1	100
パネル2 (CHP2)					
iSeq 100	97.3	0.20	100	100	100
MiniSeq	97.1	0.17	99.0	100	100
MiSeq	96.8	0.15	100	98.6	100
パネル3 (Focus)					
iSeq 100	96.5	0.23	99.5	95.2	100
MiniSeq	96.6	0.20	98.9	92.9	100
MiSeq	96.1	0.19	99.0	94.0	100

a. 平均0.2×パーセント以上

b. BRCA1パネルはBRCA1およびBRCA2遺伝子の全エクソン領域に渡る265アンプリコンを含む。CHP2パネルは50のがん遺伝子とがん抑制遺伝子をカバーする207アンプリコンを含む。Focusパネルはマルチバリエント解析のための53遺伝子の269アンプリコンを含む。

A. NextSeq 550システムラン



B. iSeq 100ラン

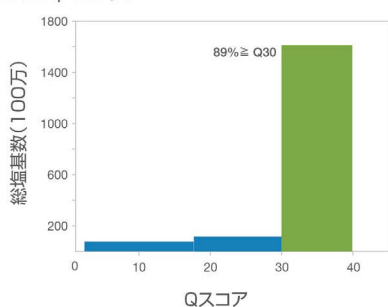


図 4: NextSeq 550およびiSeq 100システムのクオリティスコア

(A) NextSeq 550システムの2色法SBSテクノロジーは、シーケンスされた塩基の88%がQ30以上のシーケンス精度がある高品質なデータを示しました。ヒト細胞株混合サンプルをプールし、2 × 101 bp設定でランを行いました。(B) iSeq 100 システムの1色法SBSケミストリーは同等の高品質なデータを示しました。2 × 151 bpでシーケンスした微生物サンプルのプールに対して実施したメタゲノム解析の結果は、シーケンスされた塩基の89%がQ30以上を示しました。

全システムに渡る高精度なSBS

イルミナSBSケミストリーは、事実上どのカバレッジレベルでも極めて高い精度のシーケンスデータと、エラーのないリードの高い収量と、業界で最も高いQ30*を超える高品質な塩基の割合を実現しています。^{2,6,7} 他のイルミナプラットフォーム同様、iSeq 100 システムもQ30を超える高精度な塩基を高い割合で産生します (図4)。バリエント同定に対して、iSeq 100 システムで生成されたデータとMiniSeqおよびMiSeqシステムで生成されたデータと比較して評価するために、3システム全てで3つのアンプリコンパネルのランを行い、解析を行いました (表1)。Burrows-Wheeler aligner (BWA)、⁸ Piscesバリエントコーラー、⁹ およびHap.pyベンチマーク用ソフトウェア¹⁰ を使ってデータ解析を行いました。精度およびリコール率はPlatinum Genomesリファレンスデータセットを使って算出しました。¹¹ 結果では、3システムとも同等の高品質アライメント、カバレッジ、バリエント解析パラメーターが示されています。

全てのイルミナシーケンサーシステムは同じ可逆的ターミネーターによるSBSテクノロジーを使用しているため、あるイルミナシーケンサーシステムから別のシステムに不安なく移行することができます。データが4色、2色、1色法SBSで産出されても、結果はイルミナゲノムコンピューター環境のBaseSpace™ Sequence Hubまたは幅広いサード・パーティーの解析ツールで比較、解析することができます。

まとめ

イルミナSBSケミストリーは、高精度な1塩基ごとのシーケンスとゲノムの安定したカバレッジを提供します。CMOSセンサー検出と1色法SBSケミストリーを組み合わせたiSeq 100 システムは、SBSテクノロジーの最新進化形です。シンプルなケミストリーと光学技術を採用し、お手頃な価格と小さなサイズのiSeq 100 システムは、全ての研究室の日々の研究ツールとしてNGSのパワーを発揮することを可能にします。

*クオリティスコア (Qスコア) は、ベースコーリングにおけるエラー確率の予測指標です。Qスコア30 (Q30) は、高品質なデータの基準として広く認識されています。⁵

詳細はこちらから

iSeq 100 システムに関する詳細はこちらをご覧ください。
jp.illumina.com/iseq

iSeq 100 システムでの微生物またはミトコンドリアシーケンスに関する詳細は、[iSeq100システムによる微生物WGS](#)または、[iSeq100システムによるミトコンドリアDNAシーケンス](#)の Application Note をご覧ください。

パターン化フローセルに関する詳細はこちらをご覧ください。
[パターン化フローセルページ](#)

参考文献

1. Data calculations on file. Illumina, Inc., 2015.
2. Bentley DR, Balasubramanian S, Swerdlow HP, et al. [Accurate Whole Human Genome Sequencing using Reversible Terminator Chemistry](#). *Nature*. 2008;456(7218):53-59.
3. Precone V, Monaco VD, Esposito MV, et al. [Cracking the Code of Human Diseases Using Next-Generation Sequencing: Applications, Challenges, and Perspectives](#). *Biomed Res Int*. 2015;161648.
4. Shokralla S, Porter TM, Gibson JF, et al. [Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform](#). *Sci Rep*. 2015;5:9687.
5. Illumina (2011). [Quality Scores for Next-Generation Sequencing](#). Accessed November 30, 2017.
6. Ross MG, Russ C, Costello M, et al. [Characterizing and measuring bias in sequence data](#). *Gen Biol*. 2013;14:R51.
7. Nakazato T, Ohta T, Bono H. [Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive](#). *PLoS One*. 2013;22;8:e77910.
8. Li H, Durbin R. [Fast and accurate short read alignment with Burrows-Wheeler transform](#). *Bioinformatics*. 2009;25:1754-60.
9. Pisces Variant Caller. github.com/Illumina/Pisces/wiki. Accessed December 4, 2017.
10. Hap.py Benchmarking. www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/hap-py-benchmarking.html. Accessed December 4, 2017.
11. Platinum Genomes. Eberle, MA et al. [A reference data set of 5.4 million phased human variants validated by genetic inheritance from sequencing a three-generation 17-member pedigree](#). *Genome Res* 2017;27:157-164.

イルミナ株式会社

〒108-0014 東京都港区芝5-36-7 三田ベルジュビル22 階

Tel (03) 4578-2800 Fax (03) 4578-2810

jp.illumina.com

www.facebook.com/illumina

代理店

本製品の使用目的は研究に限定されます。診断での使用はできません。 販売条件： jp.illumina.com/tc

© 2017 Illumina, Inc. All rights reserved.

Illumina, BaseSpace, BeadArray, BeadXpress, cBot, CSPPro, DASL, Design Studio, GAllx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Innuity, iSelect, MiSeq, Nextera, NextSeq, NovaSeq, NuPCR, SeqMonitor, Solexa, TruSeq, TruSight, VeraCode, the pumpkin orange color, the Genetic Energy streaming bases design は、Illumina, Inc. の商標または登録商標です。

その他の会社名や商品名は、各社の商標または登録商標です。予告なしに仕様および希望販売価格を変更する場合があります。

Pub.No. 770-2013-054-B-JPN

illumina