

シングルセル生物学のマルチオーム層を包括的に解析する ResolveOMEプラットフォーム

著者

Jeff Blackinton, PhD

Director
New Product Introduction
BioSkryb Genomics

Tatiana Morozova, PhD

Senior Scientist
Research & Development
BioSkryb Genomics

Jon Zawistowski, PhD

Senior Director
Research & Development
BioSkryb Genomics

Isai Sales-Gonzalez, PhD

Computational Biologist
Bioinformatics
BioSkryb Genomics

Durga Arvapalli, PhD

Scientist I
Research & Development
BioSkryb Genomics

Swetha Velivela, PhD

Scientist I
Research & Development
BioSkryb Genomics

Jay A.A. West, PhD

President, CEO, Co-founder
BioSkryb Genomics

Bioskryb Genomics社は、シングルセルの全トランスクリプトームおよび全ゲノム増幅シーケンス解析のための統一システムを開発しました。ここでは、シングルセルのRNAとDNAコンポーネントを組み合わせた解析の統合を可能にする全体的なワークフローと必須のデータ測定基準について述べます。このシステムにより、哺乳類組織を構成する個々の細胞の転写および翻訳分子層の両方におけるゲノム変異の包括的な描写が可能になります。

製品のハイライト

- シングルセルの完全なトランスクリプトームとゲノムを解析するマルチオミクスアプローチ
- 全長転写産物を用いた全mRNAトランスクリプトーム解析
- 全ゲノム増幅 (WGA) で可能な解析例 :
 - 全ゲノムシーケンス (WGS)
 - 全エキソームシーケンス (WES)
 - ターゲットパネルによる濃縮
- オリゴ結合抗体を用いた標的タンパク質分析にも適用可能

イントロダクション

分子生物学のセントラル・ドグマは、すべての細胞のゲノムには、すべての生物学的生命に関する完全な設計書一式が含まれていると述べています。全ゲノム、トランスクリプトーム、プロテオームで構成されるオミクス層が同調して、この遺伝暗号を翻訳・解読し、細胞の運命を決定します。正常な細胞発生も病的な細胞発生も、この変容可能なプロセスの忠実性に依存しています。それらはOME、すなわちすべての部分の総和に依存しています。

このような個々の分子層やOMEの解明は、数十年にわたり大きな関心を集めており^{1,2}、シングルセルのゲノムシーケンシング³⁻⁵、トランスクリプトームシーケンシング^{6,7}、シングルセルにおける標的タンパク質検出法など、複数のオミクスベースの方法論の開発を牽引してきました⁸⁻¹⁰。

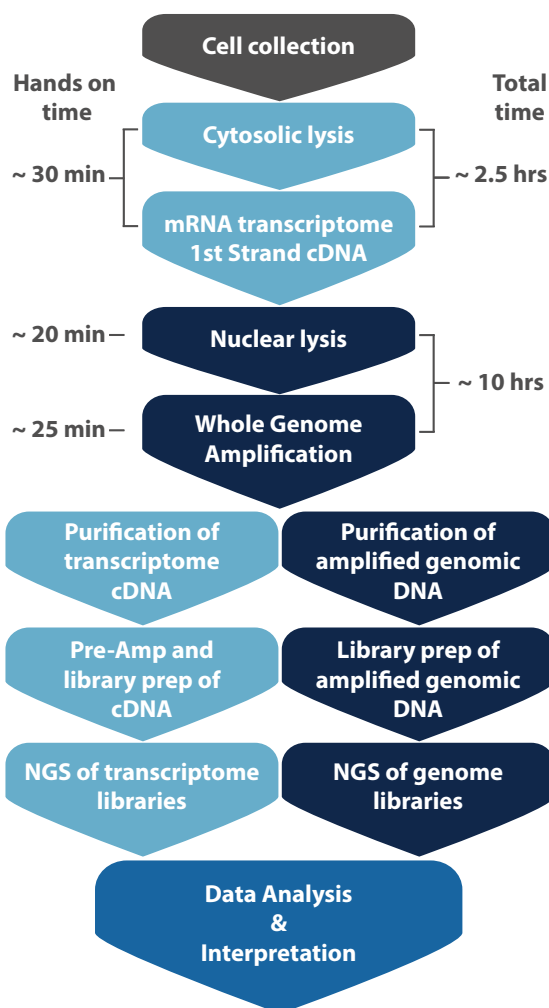


図1. ResolveOMEシステムワークフローは、1st strand cDNA合成とゲノム増幅を統合し、下流でマルチオミクスライブラリを分離してNGS解析を可能にします。※現在販売中のバージョンのキットとワークフローが異なります。

しかし、これらの方法論、特に高いカバレッジのゲノム解析を組み合わせることは、データの質の点で困難でした¹⁰。ResolveOME製品ソリューションは、この課題を解決するために開発されました。ワークフロー（図1）は、シングルセルの単離から始まり、細胞質溶解と逆転写へと続きます。これが完了すると、同じサンプルがPrimary Template-directed Amplification (PTA) 法を用いて処理され、全ゲノム増幅が行われます。ビーズを用いたトランスクリプトームcDNAプールの分離と増幅の後、トランスクリプトームとゲノム画分の両方から、下流の次世代シーケンシングと解析のための別々のライブラリが調製されます。

ResolveOMEは、PTA法の基礎の上に構築されています。MDAのような、シングルセルやピコグラム(pg)量のDNAを全ゲノム増幅するこれまでの方法では、確実なバリエーション検出に必要なゲノムカバレッジ量と均一性を得ることができませんでした³⁻⁵。PTA法¹¹の開発により、個々の細胞や少ないDNAインプットでも、高い均一性で増幅できるようになり、幅広いアプリケーションのシーケンスと最高のバリエーションコール感度が得られるようになりました。

BioSkrby Genomics社は、PTA法の性能を基に、トランスクリプトーム全体の1st strand cDNA合成を行う機能を統合しました（図2）。標準的なマイクロタイタープレートにシングルセルを単離した後、BioSkrby Resolve RTケミストリーを各細胞に適用し、細胞質溶解と各細胞からのmRNAの逆転写を行います（図2A）。各転写産物から合成されたcDNA分子は、核の溶解とそれに続く各細胞のゲノムを増幅するステップの間、サンプル中に残ります。ResolveDNA製品と同様に、各細胞のゲノムはその後、ランダムプライミングに基づくゲノム増幅に備えて変性されます（図2B）。PTA法は等温増幅と独自のターミネーションケミストリーを利用してアンプリコンのサイズを制限し、ランダムプライマーを一次テンプレートに優先的に向かわせます（図2B）。これはResolveOMEケミストリーシステムの重要な特徴であり、合成されたゲノムのアンプリコンだけでなく、最初のテンプレートスイッチに基づく1st strand cDNA合成から生成されたアンプリコンの再コピーを最小限に抑えます。

結合されたプロセスの結果、増幅されたゲノムのプールは、細胞内のcDNAの完全な多様性も含んでいます。これらの結合プールは次に、独自の磁気ビーズベースの試薬システムを用いてcDNAを分離し、

トランスクリプトームプールから増幅ゲノムDNAを除去します。ゲノムDNAとトランスクリプトームcDNAという2つの異なる画分を作製した後、最適化されたResolveOMEライブラリ調製システムを用いて、トランスクリプトームライブラリ（前増幅を含む）とゲノムライブラリのためのプロセスを用いて、サンプルをシーケンスライブラリに変換します。

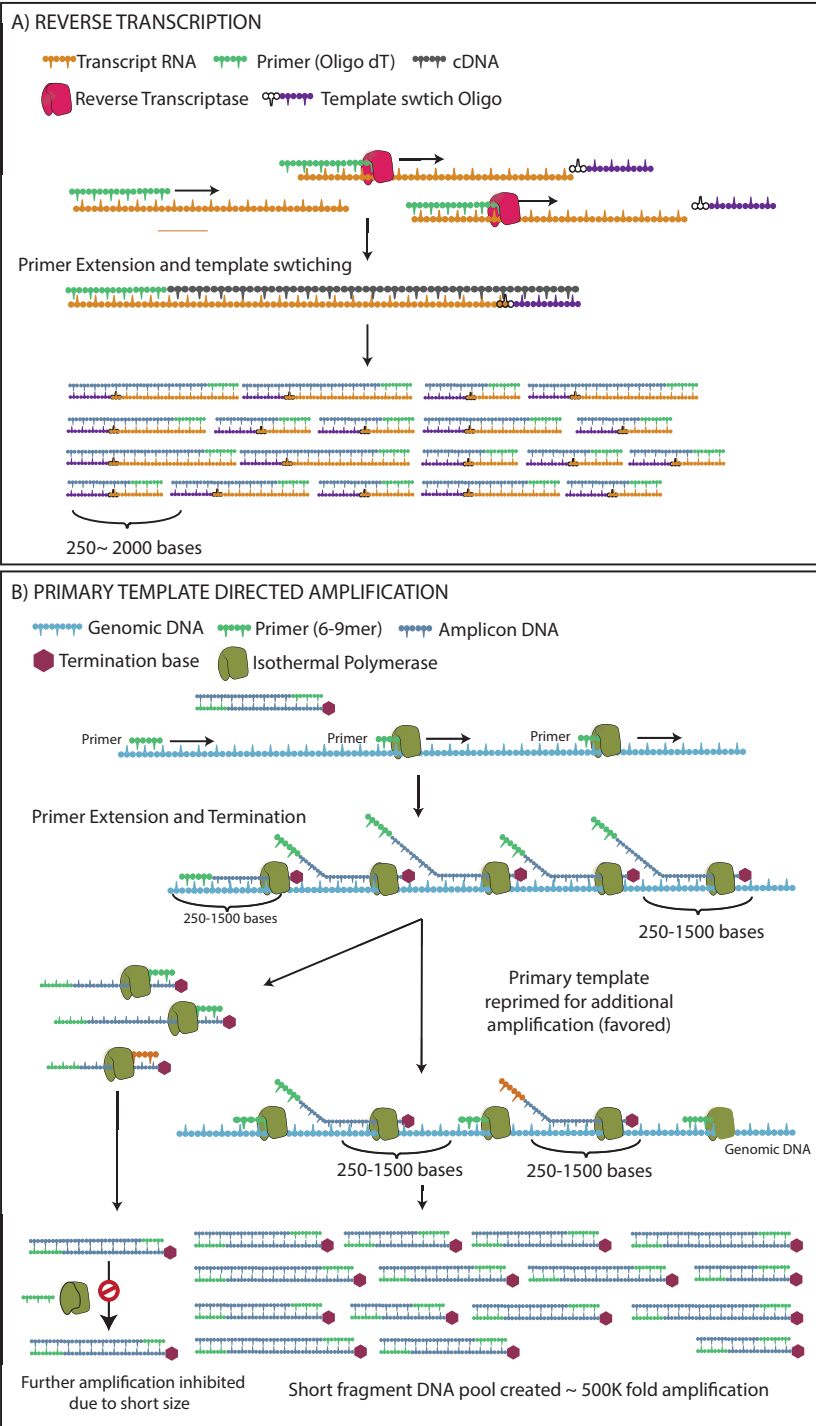


図2. ResolveOME™増幅プロセス。単離したシングルセルはまず、逆転写プロセスにより、各細胞内に存在する細胞質mRNA分子の1st strand cDNA合成を行う（A）。このプロセスが完了すると、核画分の一次テンプレート増幅が行われる（B）。その後、増幅したゲノムDNAプールから1st strand cDNA産物を濃縮し、ResolveOMEライブラリ調製システムを用いたライブラリ調製、次世代シーケンシング、BaseJumperソフトウェアによる解析を行う。

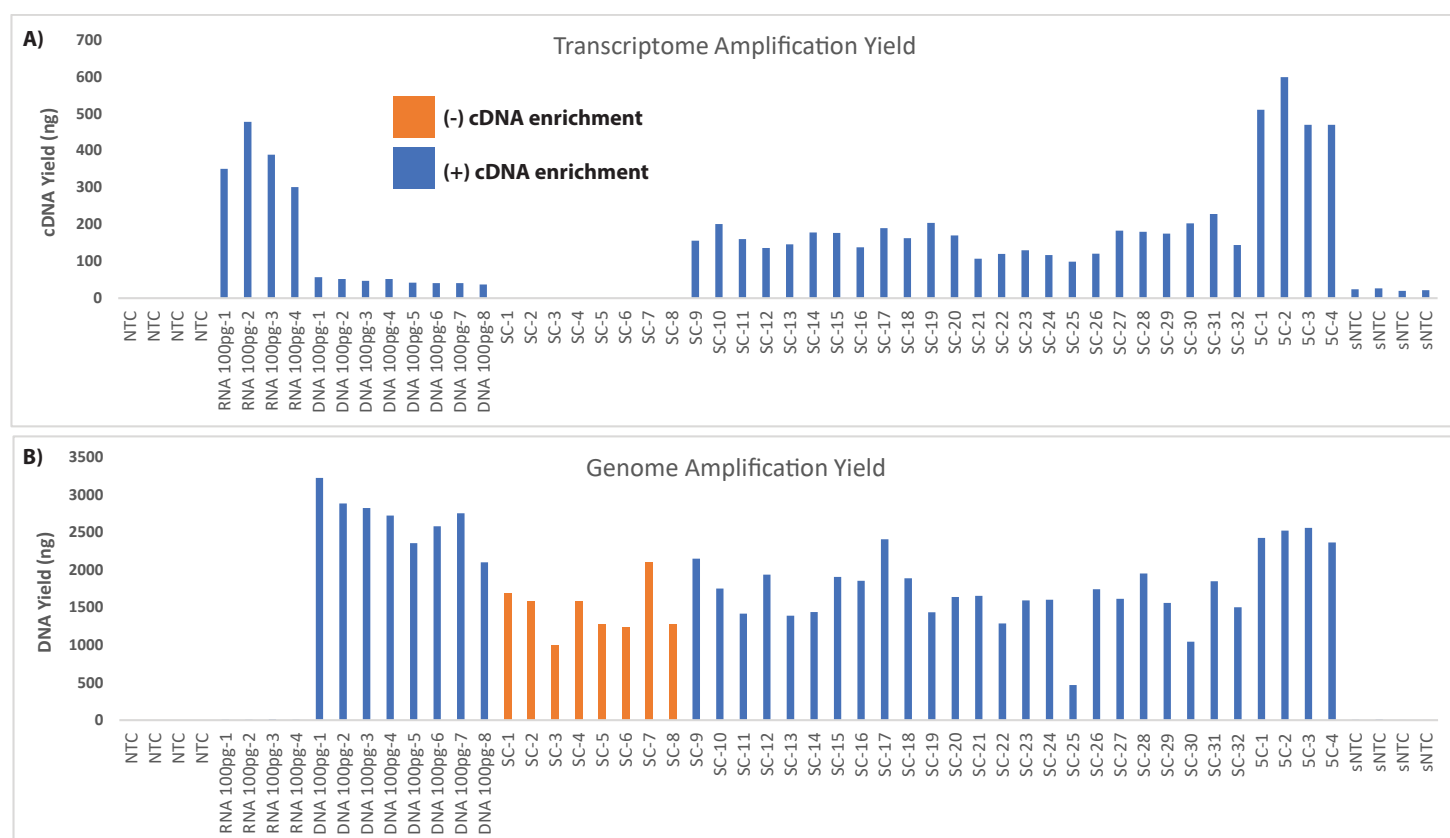


図3. 個々のGM12878細胞から得られたトランスクリプトームおよびゲノムの増幅収量。ResolveOMEケミストリーの性能を評価するために、まずトランスクリプトームとゲノムを組み合わせた増幅後のサンプルを定量した。(A)では、あらかじめ増幅したトランスクリプトームcDNAの収量を、RNAとDNAの両方、および1個の細胞、5個の細胞を含むウェルで評価した。増幅反応ではテンプレートなしコントロール (NTC) を、細胞ソーティングではFACSソーティング陰性対照 (sNTC) を同時に実施した。(B)では、増幅ゲノムDNAの収量を同じ細胞から評価した。トランスクリプトームcDNAまたは増幅ゲノムDNAの収量 (オレンジ色のバー) に対するcDNA濃縮の効果を調べるために比較を行った。その結果、トランスクリプトームcDNA量を十分得るためには、濃縮プロセスを行うことが重要であることがわかったが、増幅ゲノムDNAの収量には差が認められなかった。

このシステム全体では、全長mRNA転写産物由来のcDNA分子の解析だけでなく、各細胞の全ゲノムの解析が可能である。WGS解析は下流の次世代シーケンシングアプリケーションとして望ましいですが、ゲノム画分はWES用に濃縮することも、標的パネル濃縮に使用することもできます^{12,13}。様々な濃縮ストラテジーは、ResolveDNA全ゲノム増幅キットや他の市販ゲノム濃縮製品との統合に採用されてきたプロセスと同様に使用可能です。

実験方法および結果

増幅の効率を実証するために、まずDNA収量 (図3) と増幅後の断片サイズ (データは示さず) を分析しました。ゲノム/トランスクリプトーム結合ワークフローのパフォーマンスを調べるため、まず1000 Genomes細胞株GM12878を利用しました。この細胞株は活性のあるmRNA転写産物は少ないですが、この高度に定義されたゲノムを使用することにより、アレルのバランス、一塩基バリエーション (SNV) の精度および感度を評価することができます。ResolveOMEプロセスにより、トランスクリプトームcDNAと増幅ゲノムDNA (図3) の別々のプールを作製できることで、下流の次世代シーケンシングのためのライブラリを作製することがわかりました。シングルセルからのトランスクリプトームcDNAの平均的な収量は、前増幅後~150 ngであったのに対し、同じシングルセルから通常~1500 ngの収量の増幅ゲノムDNAが得られることがわかりま

した。精製されたDNAを用いた反応では、逆転写活性がある程度認められました (図3A、DNA 100 pg)。これらの反応では、核溶解前に細胞質で逆転写を行うシングルセルでの反応とは異なり、逆転写ステップ中にDNAが利用可能となります。この増幅産物の生成は逆転写酵素の乱雑さの結果であり、逆転写酵素は鋳型となるmRNA分子がなくてもDNAを鋳型として使用することができます。ResolveOMEのシングルセルの逆転写プロセスでは、逆転写ステップの前に核ゲノムDNAが変性しないので、このようなことはあまり起こりません。このケミストリーにおける重要な特徴は、PTA反応中にバレルRNAのトランスクリプトームcDNAを増幅する量が限られていることです (図3B)。これは、断片サイズの小さいテンプレートの再コピーが制限されているPTA法のメカニズムによるものです。

ResolveOMEプロセスにおける重要な特徴は、ゲノム増幅プロセスの完了後に行われる、ビーズを用いたトランスクリプトーム由来cDNAの濃縮です。この濃縮プロセスは、十分な収量のトランスクリプトーム由来cDNAを得るために重要である一方 (図3A)、増幅されたゲノムDNAの収量 (図3B) の損失を引き起こさないことがわかりました。シングルセルからこれら2種類のオミクスプールを増幅する場合、cDNAや増幅ゲノムDNAの収量にある程度のばらつきが生じることが予想されましたが、収量の一貫性が高いことが示され、収量が低かった細胞は1つだけ (図3B-SC-25) であったことから、FACS

の脱落やユーザーエラーの可能性が示唆されました。全体として、ResolveOMEシステムは、下流の解析のための次世代シーケンシングライブラリを調製するのに十分な量の増幅cDNAとゲノムDNAのプールを作製することができます。統合されたワークフローを使用して、同じシングルセルからのトランスクリプトームcDNAと増幅ゲノムDNA産物をライブラリ調製に使用し、リファレンスGM12878ゲノムを使用した様々な細胞タイプのパフォーマンス、発現遺伝子の検出、アレルバランスの算出、およびSNV精度および感度を比較しました。

トランスクリプトームシーケンシングによる遺伝子検出

トランスクリプトーム解析をゲノム解析と組み合わせることができるかどうかは、これら2つの分子層のデータの質が高いかどうかにかかっています。重要な指標は、シングルセル内で検出されたユニークな遺伝子の数です。この評価を作成するために、ResolveOMEケミストリーを用いて細胞タイプ間で検出された遺伝子を比較するバイオインフォマティック・アルゴリズムであるSALMONを使用しました。GM12878株をベースとして、検出された遺伝子数をいくつかの細胞タイプと比較しました（図4）。ResolveOMEケミストリーを用いたGM12878細胞で検出された遺伝子数は2500から約4500でした。

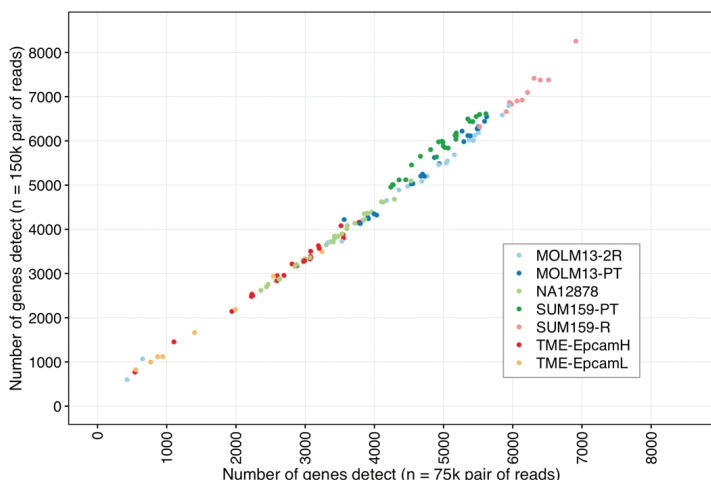


図4. SALMONによって計測された検出された遺伝子。ResolveOMEによるユニークな遺伝子の検出をさまざまな細胞タイプと条件下で測定し、シーケンスパフォーマンスを判定した。

別の細胞タイプ、特に薬剤耐性を誘導された細胞タイプでは、検出されたユニークな遺伝子の数は～4000-7000の範囲でした。この結果を、凍結保存され培養されていない臨床乳がんサンプルと比較したところ、遺伝子検出の範囲は500-4000個でした。これらのデータを総合すると、トランスクリプトームとゲノムを統合しても、データの質を犠牲にすることはないことが示唆されました。むしろ、全長転写産物の解析は、壊れやすく貴重な臨床サンプルを用いた場合でも、特定の転写産物内のスプライスバリエーションや遺伝子融合を明確にすることを可能にしながら、高い感度とダイナミックなレンジを提供します。

ゲノムシーケンシングの精度と感度

PTA法をテンプレートスイッチングによるトランスクリプトームcDNA調製と統合することで、多くの下流解析オプションが可能になります。これらの下流バイオインフォマティクス・ワークフローの中には、コピー数変異（CNV）と、染色体レベルで欠失または重複した発現遺伝子のセットを組み合わせることも含まれます。基本的なレベルでゲノムの構造を明らかにしながら細胞の表現型と状態を定義する能力は、様々な形態の病態にとって重要です。その他のアプリケーションとしては構造的に修飾されたmRNA分子と組み合わせ、病態の発がん性SNVドライバーの解析が考えられます。これらの複合的な手法による可能性は豊富にあります。しかし、相乗的な解析を行うためには、データの質が高くなければなりません。

ゲノムシーケンスにおいて最高品質のデータを示す指標はSNVの精度と感度に関するものです。これは、細胞の大規模集団（ここではバルクと呼ぶ）のシーケンスにも、シングルセルのゲノムにも当てはまります。個々の細胞には通常1つのゲノム（2倍体）、約6.6 pgのDNAが含まれることを考えると、このゲノム材料を増幅してもアレルのバランスを保たなければなりません。これは、増幅過程の均一性あるいは均等性の尺度です。PTA法は、既存の全ゲノム増幅ケミストリーの中で最も高いアレルバランスを示します¹¹。

ResolveOMEシステムも同じ要件を満たすよう開発されました。シングルセルのアレルバランス（図5）を測定することは、増幅バイアスを判定する最良の方法です。同じ細胞（図3）を用いてライブラリを調製し、品質チェックを行い、高深度シーケンスを行いました。適切なゲノムカバレッジと均一性を確認するため、まず最低2Mの総リード数で各ライブラリをシーケンスしPreseqカウントを確立しました。これは、正確な解析に通常必要とされる深度で、高い精度と感度でバリエーションをコールする能力を予測するライブラリの複

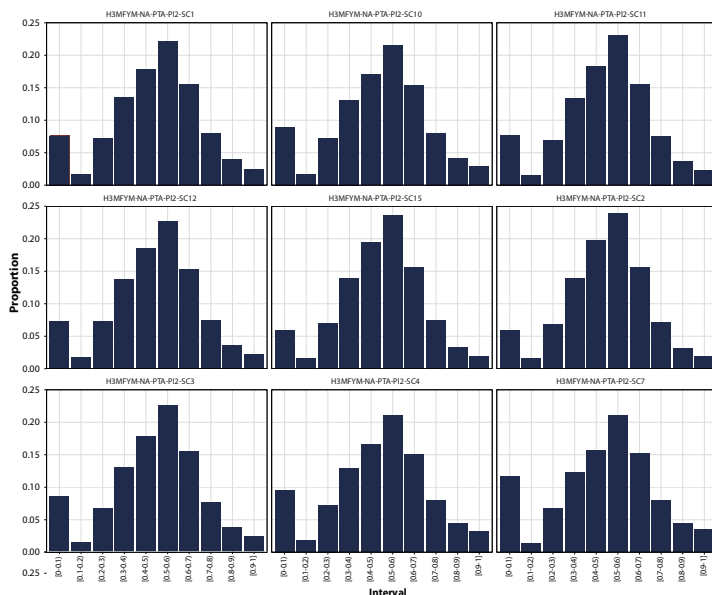


図5. ヘテロ接合性部位のアレルバランス。全ゲノムシーケンスにおいて、少なくとも2xのカバレッジを持つalternate allele頻度部位を用いて測定した。

雑さを推定するための値です¹⁴。Preseqカウントが3.5 E9以上のライブラリのみを、全深度でのシーケンスに使用しました。WGSライブラリはNovaSeq 6000 S4フローセルで2×150 bpでシーケンスしました。

シングルセルや低インプットサンプルにおけるSNV検出の精度と感度を左右する重要な要因は、ゲノム内のヘテロ接合部位のアレルバランスを厳密に維持できることです。ResolveOMEで増幅されたシングルセルサンプルのゲノムは、アレルバランスが高く（図5）、すべてのサンプルでバランスが維持されていました。このデータはResolveOME WGAプロセスとResolveOMEライブラリ調製システムを統合することで、単一のインタクトなゲノムを持つサンプルから得られたアレル頻度を忠実に再現していることを示しています。

最終的に、全ゲノム増幅ライブラリのパフォーマンスは、増幅されるゲノムをどれだけ忠実に表現できるかによって決定されます。そこで、Genome in a Bottleコンソーシアムのv3.3.2 truth setでBaseSpace Variant Calling Assessment Toolを用いてGM12878のリファレンスアレルをPTAライブラリで評価しました。

重要な点は、ヘテロ接合アレル間のバリエーション検出ではアレルバイアスは限定的であり（図5）、SNVを高精度かつ高感度で検出したことです（図6）。精度と感度を解釈するため、WGSで解析したResolveOMEシングルセルサンプルを、ResolveDNAワークフローを使用した1、3、5個のGM12878細胞、および50ngと100ngのGM12878精製ゲノムDNAと比較しました。ResolveOMEのシングルセルWGSサンプルは、全体的に91%以上のアレル検出感度を示し（ただし、意図的に含めたPreseq値が低かった1つの細胞を除く。図6A）、シングルセルサンプルで一番感度が低かったものでも91%でした。ResolveOMEシステムで処理したシングルセルの精度と感度は、ResolveDNAで処理し、Illumina 社DNA調製キット（ゲノムのみ）を用いて調製された同じ細胞のライブラリに匹敵することがわかりました。驚くことではありませんが、細胞を1つ以上含むサンプル（3細胞および5細胞）、および10細胞および20細胞相当（50 pgおよび100 pg）のゲノムDNAを使用したサンプルでは、精度と感度がわずかに高い結果となりました。ResolveOME増幅およびライブラリ調製システムを用いて調製したシングルセルライブラリは、ResolveDNAおよびIllumina社DNA調製キットを用いてライブラリ調製したものと同等の性能を示しました。

これらのデータを総合すると、均一な増幅とライゲーションベースのライブラリ調製の組み合わせにより、ResolveOMEワークフローの1st strand cDNAトランスクリプトームと組み合わせた場合でも、卓越したゲノム回収とアレルバランスが可能になることが実証されました。これにより、個々の細胞内の単一ゲノムの分解能で、高精度で高感度なSNVの検出が可能になります。

ResolveOMEケミストリーの開発により、新しいクラスのシングルセル解析が可能になりました。高分子の動態を包括的に捉えることで、細胞集団内のクローン不均一性に影響を与える因子をさらに定義し、明確にすることができます。これは正常な発生過程においても、病態の発展においても同様です。このケミストリーのシステムはまた、翻訳された細胞表面マーカータンパク質の発現をプローブするCITE-seq⁹のような新しいモダリティの追加することも可能な、かなりの柔軟性を維持しています。タンパク質の状態は、BioSkrby BaseJumper分析プラットフォームを使って、ゲノムベースの変異プロファイリングにより明らかにすることができます。BaseJumperは、精選されたデータベース（ClinVar、COSMIC）を用いて、ゲノム中の全ての各タンパク質の構造と機能に関して推論するために、コード変化のグローバルなアノテーションを可能にします。このデータはさらに、タンパク質活性の機能的変化や、改変のリスクプロファイルを予測するために使用することができます。これは、機能的な変化の可能性が高いかどうか、また、非同義変化が細胞機能にとって許容範囲か、中立か、あるいは有害かどうかを判断するために、確立されたアルゴリズム（SIFT、PROVEAN、FATHOMなど）を用いて達成されます。ResolveOMEは、個々の細胞のマルチオミックススペクトルの研究を可能にします。

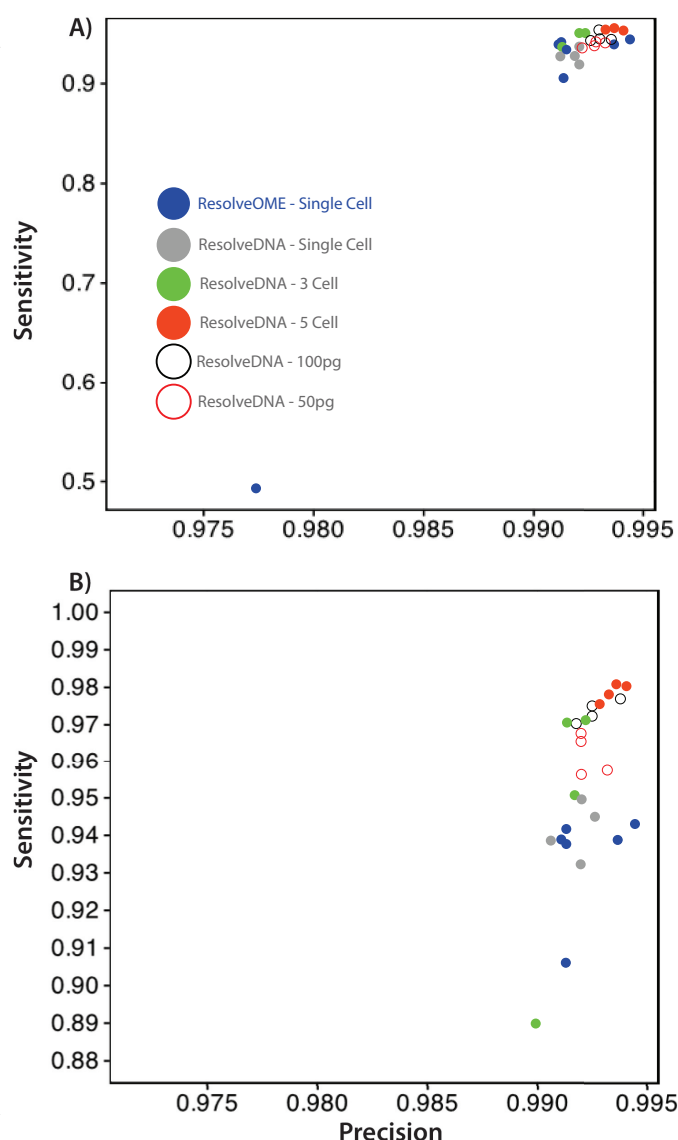


図6. 一塩基変異検出の精度と感度。反応あたり1つの細胞を含むサンプルで計測を行った。WGSのデータは、精度は99%以上、感度は91%以上で、6細胞のうち5細胞の感度は93%以上を示した。

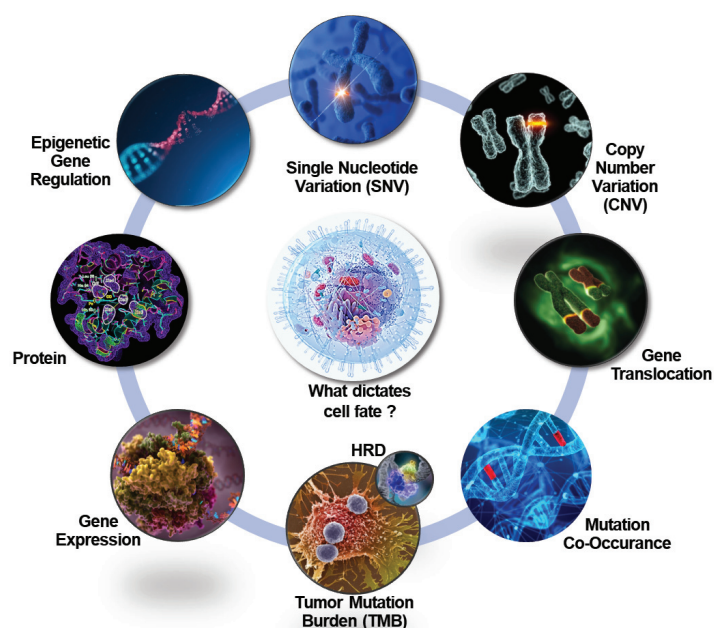


図7. 細胞の運命に影響を与える分子要因。SNV、CNV、転座、突然変異の共起、突然変異率、さらに遺伝子の発現と制御、タンパク質レベルと機能状態など、多くの要因が細胞の運命に寄与している。

要約

ResolveOMEシステムの開発により、同一細胞からトランスクリプトームとゲノムを並行して包括的に解析することが可能になりました。一塩基レベルの高分解能で精度の高いゲノム解析と、全長mRNAトランスクリプトームの包括的な解析を組み合わせて、個々の細胞内および細胞間におけるこれらのオミクス層の相互作用を理解することができます。現在のところ、マルチオミクスの理解は、同じシングルセル内のこれらの高分子層を完全に解析する能力によって制限されています。従来のマルチオミクス手法は、ゲノムのごく一部の解読、発現遺伝子の一部分の末端カウント、あるいは限られたタンパク質集団の解析に限られています。ResolveOMEシステムは、個々の細胞の最高分解能でのマルチオームビューを提供します。この方法は、シングルセル内で発現した全長転写産物を検出するバイアスのないアプローチを可能にし、同一細胞の完全ゲノムまたは標的ゲノム（例えばエクソーム）の解析を可能にします。この柔軟性を利用して、ユーザーは、発現、CNV、SNVなどの解析ベクターのサブセットを選択し、新しいバイオマーカーや潜在的な治療標的の新規発見が可能になります。ResolveOMEシステムは、腫瘍学、神経学、免疫学、生殖医療、毒物学、心臓研究など、幅広い分野での発見を可能にする、新しいレベルのマルチオミクスを可能にします。

参考文献:

1. Yu, K.H. and M. Snyder, Omics Profiling in Precision Oncology. Mol Cell Proteomics, 2016. 15(8): p. 2525-36.
2. Liu, J., et al., Applications of Single-Cell Omics in Tumor Immunology. Front Immunol, 2021. 12: p. 697412.
3. Chen, M., et al., Comparison of multiple displacement amplification (MDA) and multiple annealing and looping-based amplification cycles (MALBAC) in single-cell sequencing. PLoS One, 2014. 9(12): p. e114520.
4. Chen, C., et al., Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). Science, 2017. 356(6334): p. 189-194.
5. Dean, F.B., et al., Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A, 2002. 99(8): p. 5261-6.
6. Pollen, A.A., et al., Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat Biotechnol, 2014. 32(10): p. 1053-8.
7. Nowakowski, T.J., et al., Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science, 2017. 358(6368): p. 1318-1323.
8. Swanson, E., et al., Simultaneous trimodal single-cell measurement of transcripts, epitopes, and chromatin accessibility using TEA-seq. Elife, 2021. 10.
9. Stoeckius, M., et al., Simultaneous epitope and transcriptome measurement in single cells. Nat Methods, 2017. 14(9): p. 865-868.
10. Macaulay, I.C., et al., Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq. Nat Protoc, 2016. 11(11): p. 2081-103.
11. Gonzalez-Pena, V., et al., Accurate genomic variant detection in single cells with primary template-directed amplification. Proc Natl Acad Sci U S A, 2021. 118(24).
12. Zawistowski, J., et al., ResolveDNA Integration with Illumina DNA Prep and DNA Prep with Enrichment to Enable Single-cell Genomics. 2022, BioSkrby Genomics. p. 1-5.
13. Zawistowski, J., et al., Unprecedented Whole Exome Coverage Uniformity using ResolveDNA WGA and Twist Human Core Exome Panel, in www.bioskrby.com, B. Genomics, Editor. 2022.
14. Daley, T. and A.D. Smith, Modeling genome coverage in single-cell sequencing. Bioinformatics, 2014. 30(22): p. 3159-65.



お問合せ:

プライムテック株式会社

www.primetech.co.jp

ライフサイエンス事業部 バイオ試薬ソリューション部

東京都文京区小石川1-3-25 小石川大田ビル2F

Phone: 03-3816-0851 (代表) Fax: 03-3814-5080

E-mail: reagents@primetech.co.jp

[ウェブサイトはこちら](#)



BioSkrby

Published by: GENOMICS

2810 Meridian Parkway, Suite 110

Durham, NC 27713

www.bioskrby.com

All data on file.

For Research Use Only. Not for use in diagnostic procedures.

BIOSKRYB, RESOLVEDNA and RESOLVEOME are trademarks of BioSkrby, Inc.

All other product names and trademarks are the property of their respective owners.