

myBaits® Custom RNA-Seqを用いた cDNA RenSeqでのNLR遺伝子リシーケンシング と発現測定

概要

植物病原体防御システムの重要なコンポーネントに、ヌクレオチド結合およびロイシンリッチリピート（NLR）ドメインを持つ細胞内免疫受容体をコードする病害抵抗性遺伝子があります。抵抗性遺伝子転写産物の濃縮シーケンシング（cDNA RenSeq）は、さまざまな組織や遺伝子型で低発現しているNLR遺伝子を発見し、定量化する費用対効果の高い方法です。このアプリケーションノートでは、myBaits® RNA-Seqキャプチャシステムを使用し、様々なコムギの組織で発見したNLR遺伝子転写産物を同定および定量するcDNA RenSeqの能力を示します。このシステムがシンプルなワークフローを使用していること、正確・効率的に、そしてわずかなコストで、トランスクリプトーム全体のRNA-Seqと同等のNLR遺伝子情報を一貫して取得できることをここに示します。

なぜcDNA RenSeqを行うのか

多くのNLR遺伝子は、選択の多様化が著しいため、ユビキタスに配列の有無の多型や配列と構造の変化を伴った極度な種内多様性を示します。そのため、計算的に予測された遺伝子モデルから完全なNLR遺伝子レパートリーを取得することは不可能です[1, 2]。ゲノム RenSeqはNLR遺伝子配列の取得に焦点を当てていますが、cDNA RenSeqは発見したNLR転写産物をキャプチャし、発見していないパラログを回避することによって、機能するNLR遺伝子のより良い全体像を取得できます。さらにcDNA RenSeqは非常に発現量の少ないNLR遺伝子の発見を可能にし、自動遺伝子予測ソフトウェアによってアノテーションされていない、または誤ってアノテーションされた新規転写物およびまれなアイソフォームを特定することができます[3]。cDNA RenSeqは、ディープトランスクリプトームワイドRNA-Seqの費用対効果の高い代替手段として、高解像度の定量的データセットを生成して、NLR遺伝子の発現量の変化を評価することもできます。これらの利点により、cDNA RenSeqはNLR遺伝子の多様性を研究するための有用なツールになるだけでなく、NLR遺伝子発現のダイナミクスを研究するための不可欠な非常に効率的なアプローチにもなります。

材料と方法

myBaits® Custom RNA-Seqを用いたcDNA RenSeqの精度と効率を実証するために、8つのコムギ連(Triticeae)種に存在する予測されたNLR遺伝子から設計された60,000のBaitsを含むカスタムTriticeae RenSeq Bait-libraryを使用して生成されたデータを分析しました (Steuernagel et al. 2016 [4]を参照)。myBaits® RNA-Seqを用いた cDNA RenSeqワークフローを図1Aに示しました。概要としては、さまざまな成長段階および組織から抽出されたコムギのRNAを用いて、製造元のプロトコルに従って、KAPA mRNA-Seq Kitを使用し、cDNAライブラリを調製しました。ターゲットエンリッチメントはmyBaits標準プロトコルバージョン3に従って実行し、濃縮した、および濃縮しなかった両方のcDNAライブラリをイルミナ社HiSeq 2500でシーケンスし、図1Bに示すようなバイオインフォマティックパイプラインで下流の解析を実行しました。

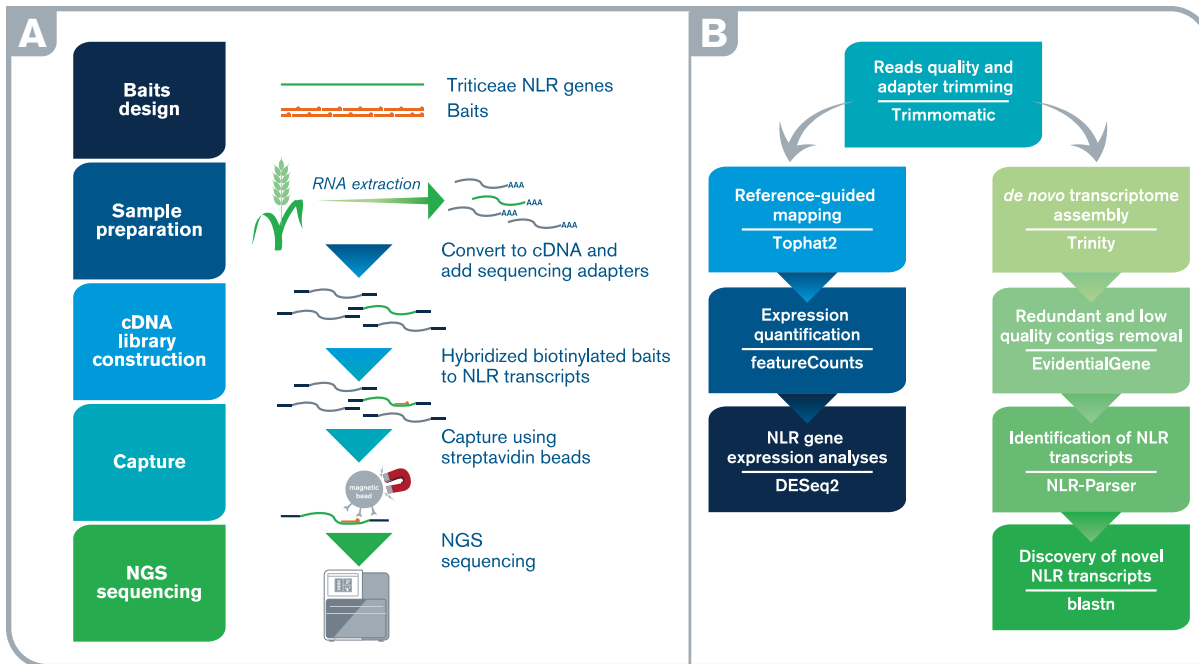


図1. シンプルで簡単な実行。myBaits® Custom RNA-Seqを用いたcDNA RenSeqワークフロー(A)とバイオインフォマティクスパイプライン(B)の概略図。

結果/考察

myBaits®を用いたcDNA RenSeqの高い濃縮効率

濃縮効率を評価するために、Steuernagelら2018[5]で報告されたデータを分析しました。ここでは、若い葉のサンプルからのcDNA RenSeqライブラリを、対応する非濃縮RNA-Seqライブラリ（ここではDeep RNASeqと呼びます）と比較しました。サンプルあたり約100GbpのDeep RNA-Seqでは、NLR遺伝子にアラインされたリードは0.16~0.31%しか得られませんでした。cDNA RenSeqデータでは、47.46~55.26%のアラインされたリードが得られました（図2）。これは約222倍の濃縮率であり、Deep RNA-Seqのシーケンス量のわずか3%を使用することで、NLRカバレッジ深度の平均4.3倍の改善につながりました（図4A）。

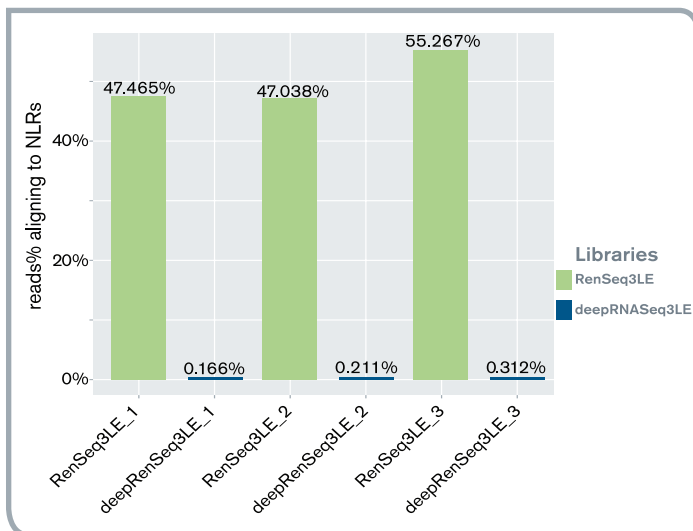
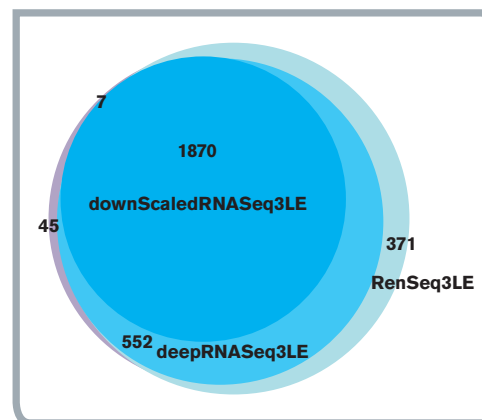


図2. 桁違いのターゲットエンリッチメント。cDNA RenSeqおよびRNA-SeqライブラリにおけるNLR遺伝子にアラインするリードの割合。

cDNA RenSeqの利点をわかりやすく説明するためにDeep RNA-SeqデータをcDNA RenSeqと同じシーケンス深度にダウンスケール（以下、ダウンスケールRNA-Seq）した後に得られたカバレッジと比較しました。ダウンスケール後、cDNA RenSeqで得られた転写産物あたりのリード深度の大幅な増加はさらに明確になりました。ダウンスケールされたRNA-Seq、Deep RNA-Seq、およびcDNA RenSeqデータセットから計算されたリード深度のlog2の中央値はそれぞれ0.3, 3.58, 5.67でした（図4A, D）。Deep RNA-Seqでの読み取り深度の変化を分析すると、30倍の深さでシーケンシングしても平均24倍の増加しか得られなかった一方、cDNA RenSeqへ切り替えると、平均212倍の増加が見られました（図4B）。

同等のリード深度でより多くのNLRの転写物が検出可能に



cDNA RenSeqは少量のNLR転写産物を検出でき、同時に転写産物あたりのリード深度を増やすことができます。図3に示すようにcDNA RenSeqはDeep RNA-

Seqよりも319（13%）多くNLR転写産物を検出し、ダウンスケールRNA-Seqよりも916（49%）多く検出しました。

図3. 発現量の低い転写産物を発見する。cDNA RenSeq、deep RNA-SeqおよびダウンスケールRNA-Seqライブラリで検出されたNLR転写産物の数。

ダウンスケールRNA-Seqで検出されなかった少量の転写産物について、cDNA RenSeqでは53%以上検出できるだけでなく、Deep RNA-Seqよりも2.5倍高い平均リード深度が得られました(図4C)。まとめると、NLR転写産物を取得するためにRNA-Seqの深度を増やすのではなく、cDNA RenSeqを使用する方が大幅に効率的で費用対効果が高くなります。

図4. リード深度の大幅な改善。(A) 各ライブラリで検出されたNLR転写物のlog2リード深度。(B) cDNA RenSeqへ切り替えた時とDeep RNA-Seqのlog2リード深度の変化。(C) cDNA RenSeqまたはDeep RNA-Seqで検出され、ダウンスケールRNA-Seqでは検出されなかった低発現のNLR転写産物のlog2リードデプス。N: NLRトランスクリプトの数。(D) cDNA RenSeq、Deep RNA-SeqおよびダウンスケールRNA-SeqでのNLR遺伝子座のリード深度の可視化。



発現測定の高さと再現性

濃縮されたライブラリが標準のRNA-Seqで測定された発現レベルをどの程度反映しているかを調べるために、cDNA RenSeqデータセットと、それに対応する濃縮されていないDeep RNA-SeqデータセットのNLR遺伝子のlog₂リードカウント（発現量）の相関関係を調査しました。ここでは、一連の発現レベルの転写物の相対的なカバレッジにおいて、優れた相関関係（R = 0.92; 図5A）が観察されました。

さらに、NLR遺伝子の正規化されたlog₂リード数を使用した階層的クラスター分析により、ライブラリはシーケンス法（直接対濃縮）ではなく組織タイプに基づいてクラスター化されることが示され、キャプチャがNLR遺伝子の発現量に系統的なバイアスを導入しないことが示されました（図5B）。再現性は、生物学的複製間の正規化されたNLRのlog₂読み取りカウントの相関係数を測定することによって試験を行い、高い平均ペアワイズ相関、R = 0.95を示しました。全体として、検証分析は、myBaits®を用いたcDNA RenSeqが正確で再現性があり、効率的なNLR発現プロファイリングのための強力な方法であることを示しています。

機能性および新規のNLR転写物の特定

ケーススタディとして、de novoアセンブリベースのパイプラインを使用して、複数組織のcDNA RenSeqを組み合わせることで、リファレンスゲノムにアノテーションされていない115の新規NLR関連転写産物を同定し、さらにアノテーションされた転写物と相同性の低い581 NLR関連転写産物の同定に成功しました。これにより、myBaits®を用いたcDNA RenSeqにより、計算遺伝子モデル予測ツールで検出されなかった新規のNLR転写物およびまれなアイソフォームを検出することが非常に効率的であることを実証されました。

結論

このアプリケーションノートは、機能性のNLRレパートリー取得におけるcDNA RenSeqと、定量的な遺伝子発現アプリケーションに使用できるmyBaits® Custom RNA-Seqのパワーを示しています。非常に効率的で信頼性が高く、手頃な価格のアプローチとして、cDNA RenSeqを実装すると、間違いなくアノテーションされたNLR遺伝子の数を増加させ、病害抵抗性遺伝子の機能研究が大幅に加速されるため、より複雑なゲノムを持つ農学的に重要な種に対して、より対象を絞った特異的な病害抵抗性育種戦略のための確固たる基盤が確立されます。

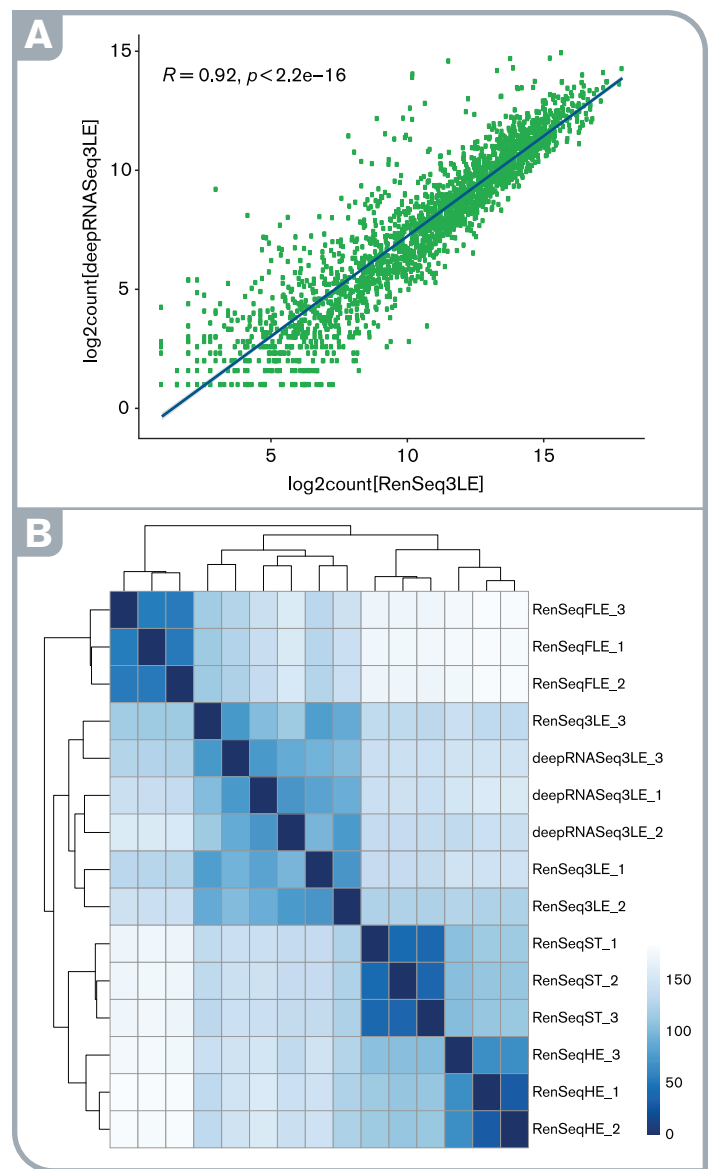


図5. cDNA RenSeqの正確性と再現性。(A) cDNA RenSeqとDeep RNA-Seq ライブラリにおけるNLR遺伝子のlog₂リード数の高い相関を示す散布図。(B) 異なる組織での遺伝子発現の階層的クラスタリング。

引用文献

1. Van de Weyer, A. L. *et al.* (2019) *Cell*.
2. Ma, Y. *et al.* (2019) *BMC Genomics*.
3. Andolfo, G. *et al.* (2014) *BMC Plant Biology*.
4. Steuernagel, B. *et al.* (2016) *Nature Biotechnology*.
5. Steuernagel, B. *et al.* (2018) *bioRxiv*.

my Baits®

NGS Hybridization Capture
www.arborbiosci.com/targeted-ngs
11/2021

4