

# cDNA RenSeqを使用したNLR遺伝子リサーチ エ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耐病性遺伝子があります。この耐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キャプチャーシステムを使用し、様々な小麦組織で発現した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に匹敵するNLR遺伝子情報を一貫して取得す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、大幅に多様化する選択により、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う極端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になります[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は発現したNLR転写産物をキャプチャし、発現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遺伝子に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注釈が付けられていない、または誤って注釈が付けられていない新規転写物およびまれなアイソフォームを特定します[3]。さらに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

cDNA RenSeqの精度と効率を実証するために、8つのTriticeae種に存在する予測される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®cDNA RenSeqワークフローを図1Aに示しました。要するに、RNAは小麦のさまざまな成長段階のさまざまな組織から抽出されました。製造元のプロトコルに従って、KAPA mRNA-Seq Kitを使用し、cDNAライブラリーを調製しました。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に従って実行されました。エンリッチメント、ノンエンリッチメントの両方のcDNAライブラリーがイルミナ社HiSeq 2500でシーケンスされ、図1Bに示すように下流の解析はバイオインフォマティックパイプラインで実行されました。

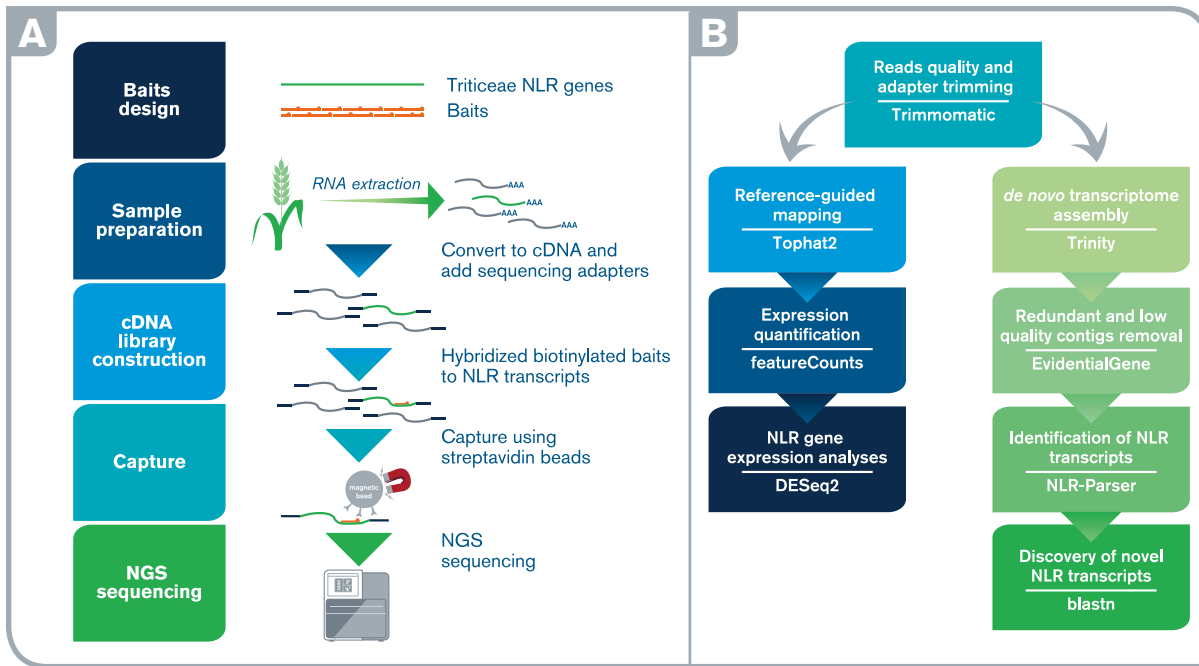


図1. シンプルで簡単な実装。

(A) myBaits@cDNA RenSeqワークフローとバイオインフォマティクスパイプラインの概略図。

## 結果/考察

### myBaits@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ライブラリーを、対応する非濃縮RNA-Seqライブラリー（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遺伝子にアラインするリードは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%のリードアラインが得られました（図2）。これは約**222倍の濃縮率**にであり、Deep RNA-Seqのシーケンスのわずか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**につながります（図4A）。

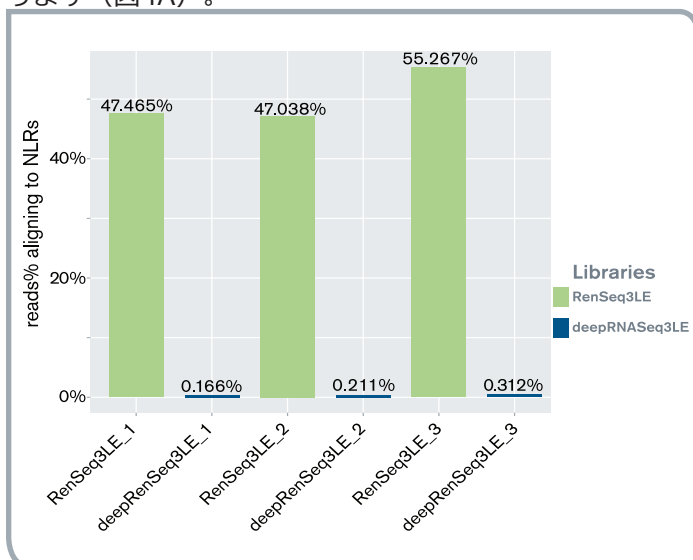


図2. 桁違いのターゲットエンリッチメント. cDNA RenSeqおよびRNA-SeqライブラリーにおけるNLR遺伝子にアラインするリードの割合

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した後に得られたカバレッジと比較しました（以下、ダウンスケールRNA-Seqと呼びます）。ダウンスケール後、cDNA RenSeqで達成される転写産物あたりの読み取り深度の大幅な増加はさらに明確です。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log2読み取り深度の中央値はそれぞれ0.3, 3.58, 5.67でした（図4A, D）。DeepRNA-Seqでは読み取り深度の変化を分析すると、30倍の深さのシーケンシングで平均24倍の増加しか得られず、cDNA RenSeqへの切り替えると、平均212倍の増加が見られます（図4B）。

### 同等の読み取り深度でより多くの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読み取り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%）多い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。

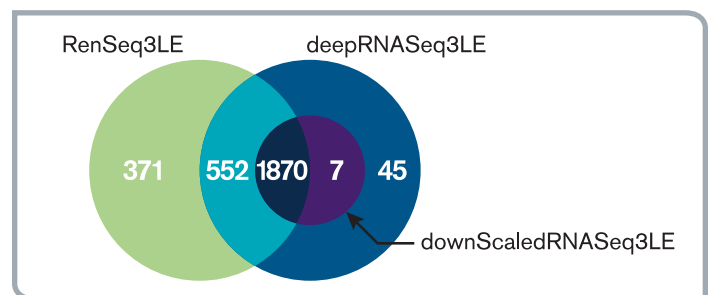
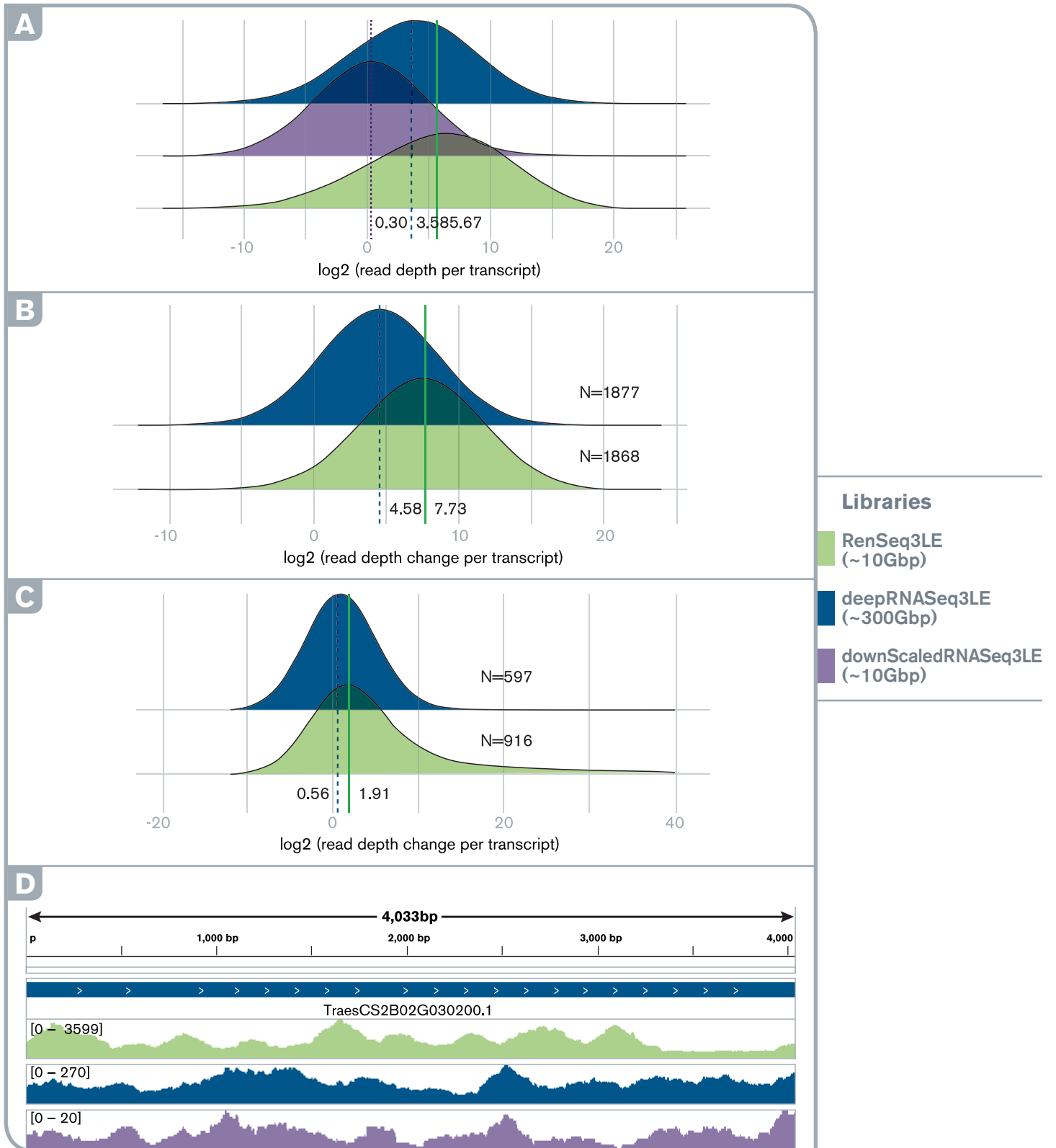


図3. 発現量の低い転写産物を発見する cDNA RenSeq、deep RNA-SeqおよびdownScaled RNA-Seqライブラリーで検出されたNLR転写産物の数。

ダウンスケールされたRNA-Seqで検出されなかった少量の転写産物を、cDNA RenSeqは53%以上の少量の転写産物を検出できるだけでなく、ディープRNA-Seqよりも2.5倍高い平均リードデプスを達成しました(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. リードデプスの大幅な改善 (A) 各ライブラリで検出されたNLR転写物のlog<sub>2</sub>リードデプス (B) cDNA RenSeqへ切り替えた時とシーケンシングデプスを増加させた時のlog<sub>2</sub>リードデプスの変化 (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log<sub>2</sub>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downScaled RNA-SeqのNLR遺伝子座のリードデプスの可視化。



## 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log<sub>2</sub> 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log<sub>2</sub>リードカウントを使用した階層的クラスター分析により、ライブラリーはシーケンス法（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、キャプチャーが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す（図5B）。再現性は、生物学的複製間の正規化されたNLRのlog<sub>2</sub>読み取りカウントの相関係数を測定することによってテストされました。これは、高い平均ペアワイズ相関 $R = 0.95$ を示します。全体として、検証分析は、myBaits@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。

## 機能的および新規の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、複数の組織cDNA RenSeqにより、参照ゲノムに注釈が付けられていない115の新規NLR関連転写産物を同定し、そして注釈付き転写物と相関性の低い581 NLR関連転写産物の同定に成功しました。myBaits@cDNA RenSeqを使用して、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/または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します。

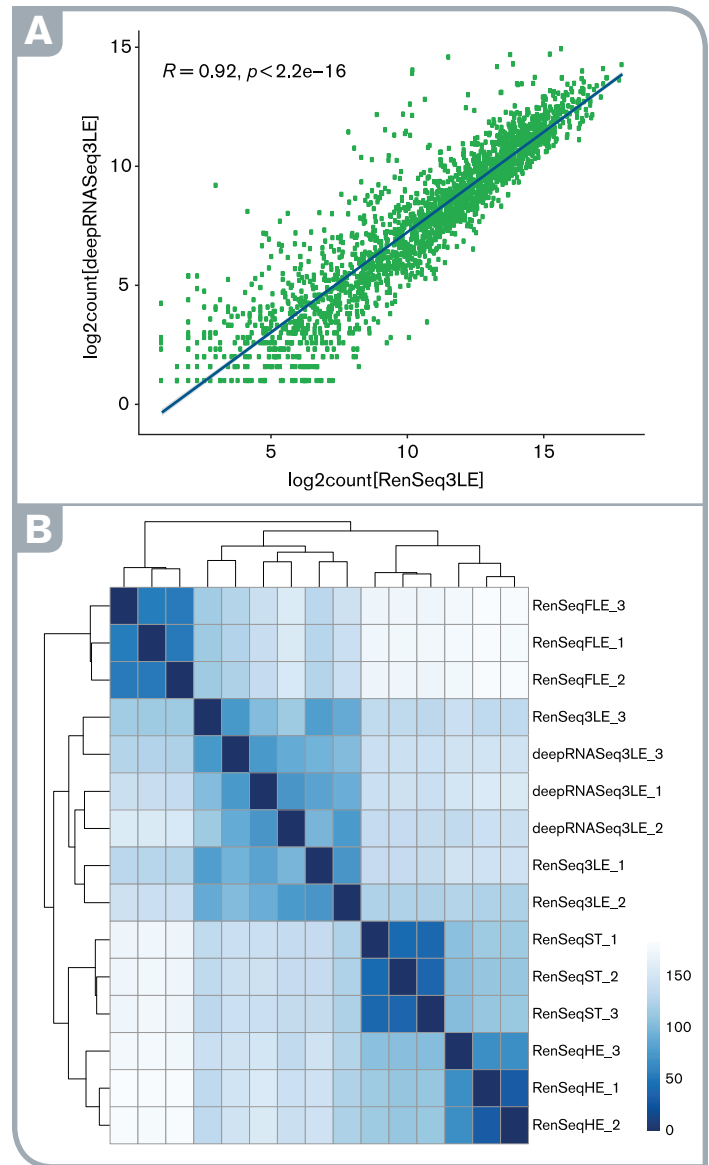


図5. cDNA RenSeqの精度と再現性 (A) cDNA RenSeqとDeep RNA-Seq ライブラリにおけるNLR遺伝子のlog<sub>2</sub>リードカウントの散布図の高い相関 (B) 異なる組織での遺伝子発現の階層的クラスタリング。

## 結論

このアプリケーションノートでは、myBaits@cDNA RenSeqの機能的なNLRレパートリーを取得するパワーと、定量的な遺伝子発現アプリケーションに使用できる能力を示しています。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、注釈付きNLR遺伝子の数が間違いなく増加し、耐病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より対象を絞った特定の耐性育種戦略の確固たる基盤が確立されます。

## REFERENCES

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*.



web [www.arborbiosci.com](http://www.arborbiosci.com)  
email [info@arborbiosci.com](mailto:info@arborbiosci.com)  
phone 1-734-998-0751  
twitter @ArborBio