

RNA-Seq

～研究に合わせたアプリケーションの選び方～

June 12, 2015



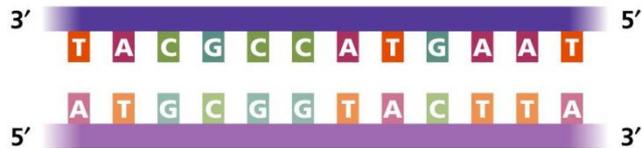
山重 りえ
イルミナ株式会社
テクニカルアプリケーションサイエンティスト

© 2012 Illumina, Inc. All rights reserved.

Illumina, illuminaDx, BaseSpace, BeadArray, BeadXpress, cBot, CSPPro, DASL, DesignStudio, Eco, GAllx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, Sentrix, SeqMonitor, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.

illumina®

DNA



- DNAは常に同一である。
- 1細胞中に常に同一の量が存在する
- すべての細胞が同じDNAを保持している（癌細胞でなければ）

Stable

RNA

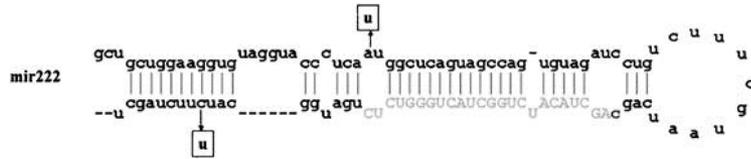


- RNAは常に変動している。
- 置かれている環境に常に影響を受ける
- 細胞間で異なる量、異なるタイプのRNAが存在する

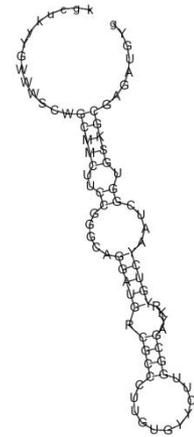
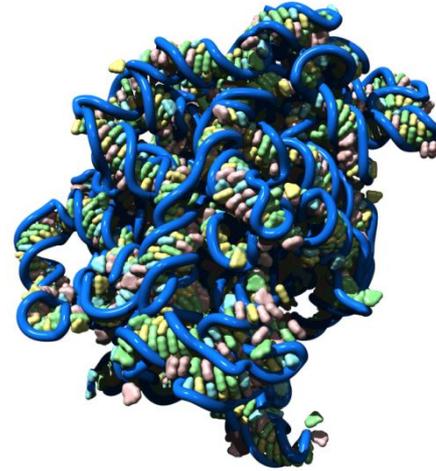
Dynamic

RNAを特徴づけるもの

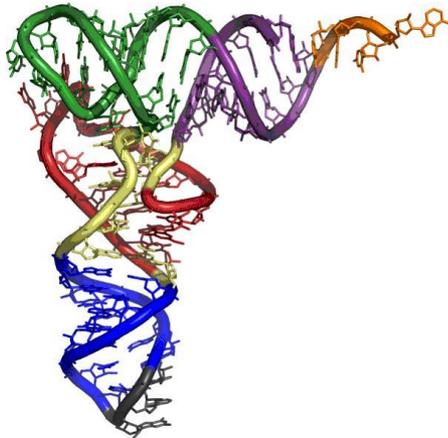
サイズ



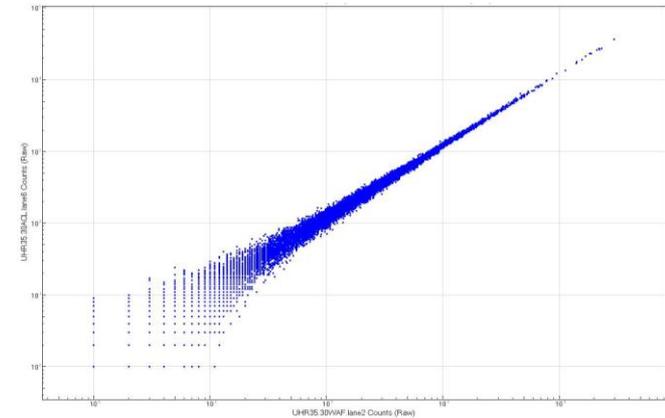
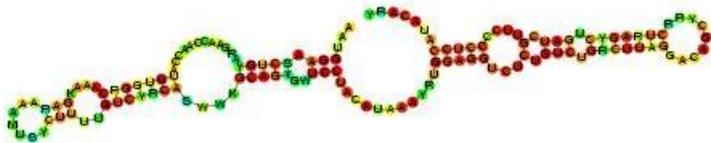
構造



安定性



コピー数



機能別のRNAの分類とシーケンス

タンパク質合成

- Messenger RNA (mRNA)
- Ribosomal RNA (rRNA)
- Transfer RNA (tRNA)
-

転写後（転写中）修飾、DNAの複製

- Small nuclear RNA (snRNA)
- Small nucleolar RNA (snoRNA)
-

発現制御

- microRNA (miRNA)
- Piwi-interacting RNA (piRNA)
- Small interfering RNA (siRNA)
- Long intergenic non-coding RNA (lincRNA)
-

機能別のRNAの分類とシーケンス

タンパク質合成

- **Messenger RNA (mRNA)**
- Ribosomal RNA (rRNA)
- Transfer RNA (tRNA)
-

転写後（転写中）修飾、DNAの複製

- Small nuclear RNA (snRNA)
- Small nucleolar RNA (snoRNA)
-

発現制御

- microRNA (miRNA)
- Piwi-interacting RNA (piRNA)
- Small interfering RNA (siRNA)
- **Long intergenic non-coding RNA (lincRNA)**
-

→ mRNA Sequencing

機能別のRNAの分類とシーケンス

タンパク質合成

- **Messenger RNA (mRNA)**
- **Ribosomal RNA (rRNA)**
- **Transfer RNA (tRNA)**
-

転写後（転写中）修飾、DNAの複製

- **Small nuclear RNA (snRNA)**
- **Small nucleolar RNA (snoRNA)***
-

発現制御

- **microRNA (miRNA)***
- **Piwi-interacting RNA (piRNA)***
- **Small interfering RNA (siRNA)***
- **Long intergenic non-coding RNA (lincRNA)**
-

→ **Total RNA Sequencing**

* Total RNAのサンプル調製では150nt以下のRNAは除かれてしまう

機能別のRNAの分類とシーケンス手法

タンパク質合成

- Messenger RNA (mRNA)
- Ribosomal RNA (rRNA)
- Transfer RNA (tRNA)
-

転写後（転写中）修飾、DNAの複製

- Small nuclear RNA (snRNA)
- Small nucleolar RNA (snoRNA)*
-

発現制御

- **microRNA (miRNA)**
- **Piwi-interacting RNA (piRNA)**
- **Small interfering RNA (siRNA)**
- Long intergenic non-coding RNA (lincRNA)
-

→ Small RNA Sequencing

RNA-Seqのアプリケーション

発現プロファイリング解析

- 遺伝子の発現比較
 - アレルごとの発現比較
-
- miRNA発現プロファイリング解析

トランスクリプトーム解析

- 転写産物のアセンブル
 - 癌細胞内での融合遺伝子の発見
 - バイオマーカーの発見
 - 新規転写産物の発見
 - RNAエディティングの解析
 - 転写開始サイト（5'-UTR領域）のマッピング
-
- ゲノムアノテーション (ESTs)
 - 選択的スプライシング
-
- Non-coding RNAsの研究
 - アンチセンスRNAの研究
 - Small RNAsの研究

一般的なRNA-Seqのワークフロー

RNA
抽出



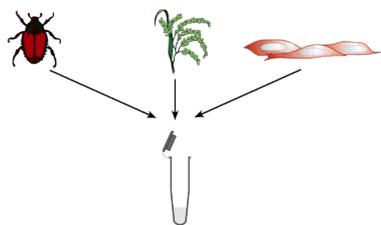
ライブラ
リー調製



シーケンス

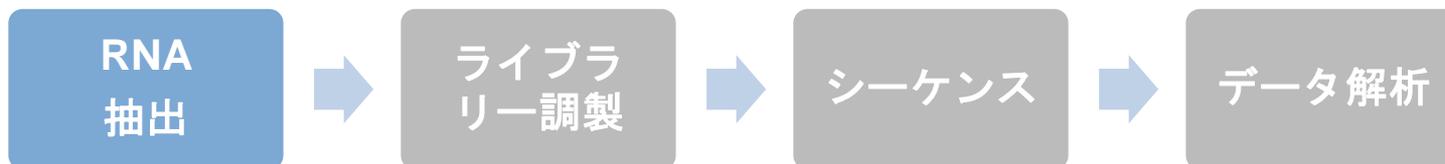


データ解析





RNA-Seqの実験デザインで考慮すべき点



RNA抽出キット

ご自分のサンプルに適したものをお選びください

mRNA & Total RNA

- **Epicenter**
 - MasterPure RNA
 - MasterPure Yeast RNA
 - MasterPure Plant RNA
- **Qiagen**
- **Ambion**
- **Trizol**
- **Others.....**

- RNA抽出キットに関しては、特に弊社から推奨のものはございません。
- 論文等をご参照の上、ご検討ください。

RNA抽出キット

ご自分のサンプルに適したものをお選びください

Small RNA

- サンプルタイプに適したものを選択（論文等参照）
 - Qiagen (miRNeasy)
 - Life technologies (mirVana)など
- 抽出時にバイアスがかからないものを選択する ← Key step!

- RNA抽出キットに関しては、特に弊社から推奨のものはございません。
- 論文等をご参照の上、ご検討ください。

RNAインプットのクオリティ

RNA-Seqを行う上で、非常に重要なポイント！

mRNA, Total RNA & Small RNA

- **TruSeq RNA v2 & Stranded mRNA**

- RIN>8、FFPEサンプルは使用不可

- **TruSeq Stranded total RNA**

- 分解の進んだRNAサンプル、FFPEサンプルなども使用可能

- **TruSeq Small RNA**

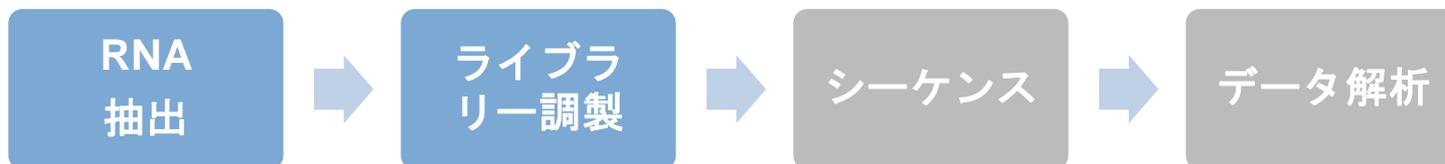
- RIN>8、FFPEサンプルは使用不可

- ▶ どのようにQCチェックを行うか？

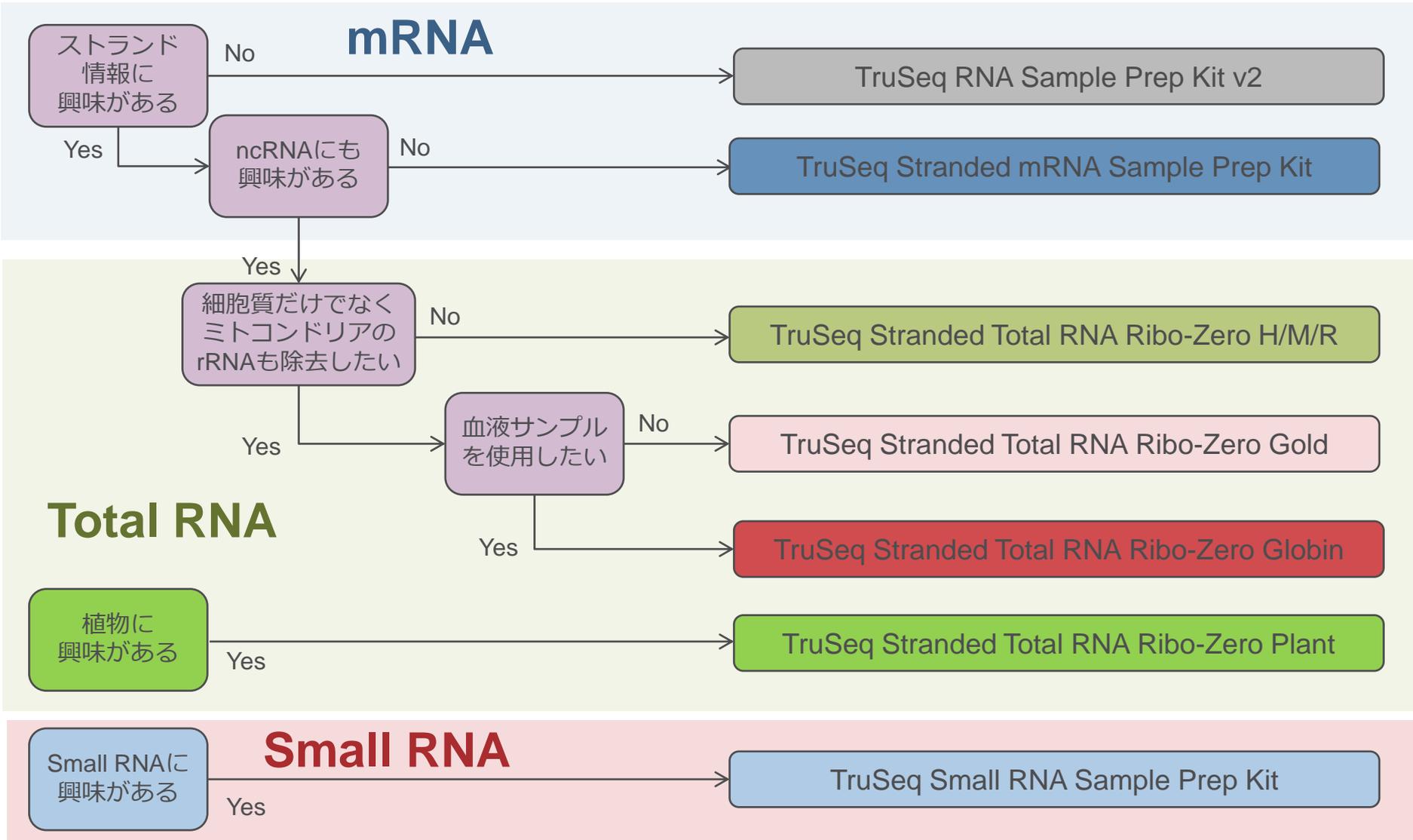
- Bioanalyzerを用いてRIN値を測定
- 変性アガロースゲル（ホルムアルデヒド 1%を含む）で電気泳動



RNA-Seqの実験デザインで考慮すべき点

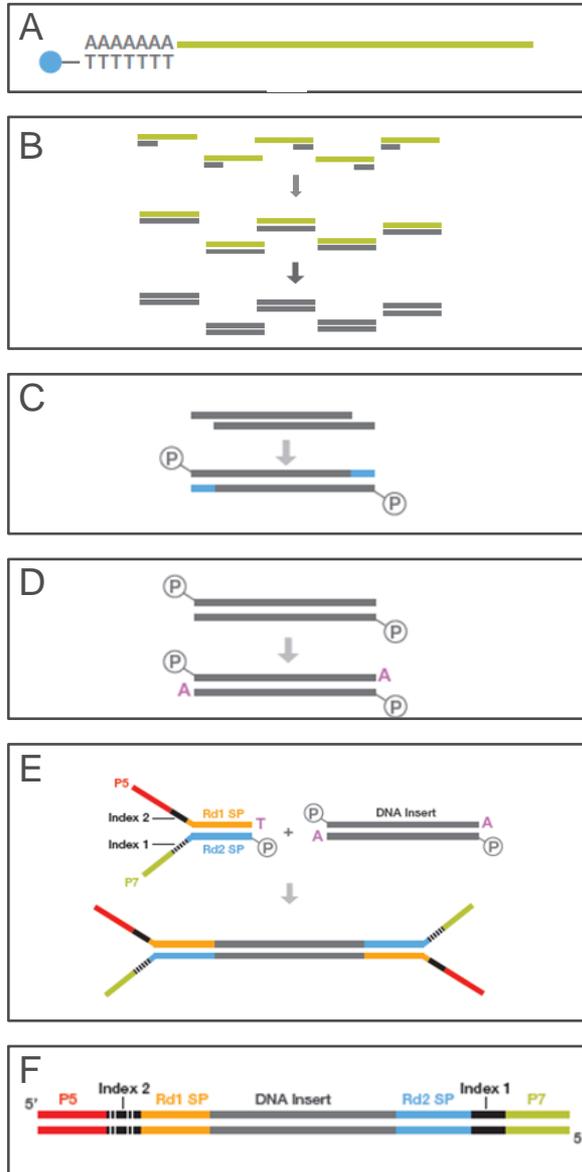


イルミナのRNAライブラリ調整キット



詳細はイルミナサポートウェビナー「RNA-Seq入門（キットの選び方、実験デザイン）」(2014/4/18)をご参照ください

TruSeq RNA v2, TruSeq Stranded mRNAワークフロー



A) Total RNAからpolyAを使いmRNAを回収

B) RNA断片化、ランダムプライマーを用いてcDNA合成

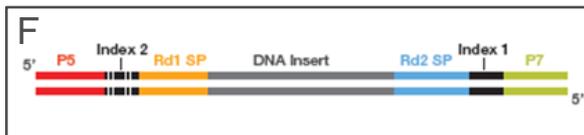
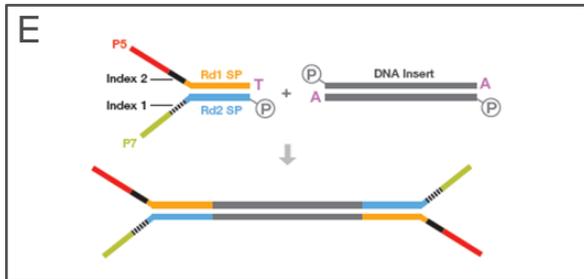
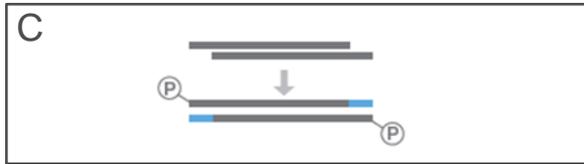
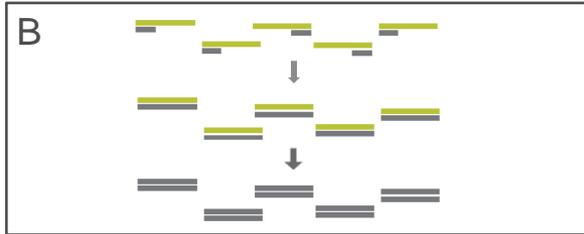
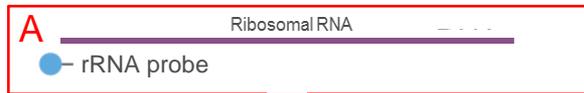
C) 末端平滑化、リン酸化

D) Aテイル付加

E) インデックス付アダプターライゲーション

F) PCR

TruSeq Stranded Total RNAワークフロー



A) Total RNAからrRNAプローブ (Ribo Zero) を用いてrRNAを除去

B) RNA断片化、ランダムプライマーを用いてcDNA合成

C) 末端平滑化、リン酸化

D) Aテイル不可

E) インデックス付アダプターライゲーション

F) PCR

Poly-A Selection vs Ribo-Zero Depletion

▶ 手法の違い:

- TruSeq RNA v2 & TruSeq Stranded mRNA: **poly A-selection 法**
- TruSeq Stranded Total RNA RNA: **Ribo-ZeroによるrRNA除去**

▶ poly A-selection法: RNAのクオリティー (RIN > 8) が重要!

- RIN<8のサンプル ->3'-側の領域にバイアスのかかったシーケンス結果になる

▶ Ribo-Zero 法: 分解の進んだRNAも使用可能!

- RIN<8のクオリティの高いRNAの使用を推奨
- 分解の進んだサンプル(RIN<3)でも3'-側の領域にバイアスはなし
- ncRNAもキャプチャー*

**poly A-selection法のサンプルに比べると～2倍のリード数が必要*

分解の進んだRNAでライブラリー調製を行う際の新しい提案： TruSeq RNA Accessライブラリー調製キット

＜これまでの手法＞

TruSeq Stranded Total RNA

- 分解が進んでいるため、polyAを利用した手法は使えない
- ランダムプライマーを用い、トータルRNAを対象に解析（rRNAは除去可能）

＜利点＞

- mRNA以外のncRNAを観察できる

＜課題＞

- 多くのリード量が必要
- mRNA以外の情報がでてくる



＜新たな提案＞

TruSeq RNA Access

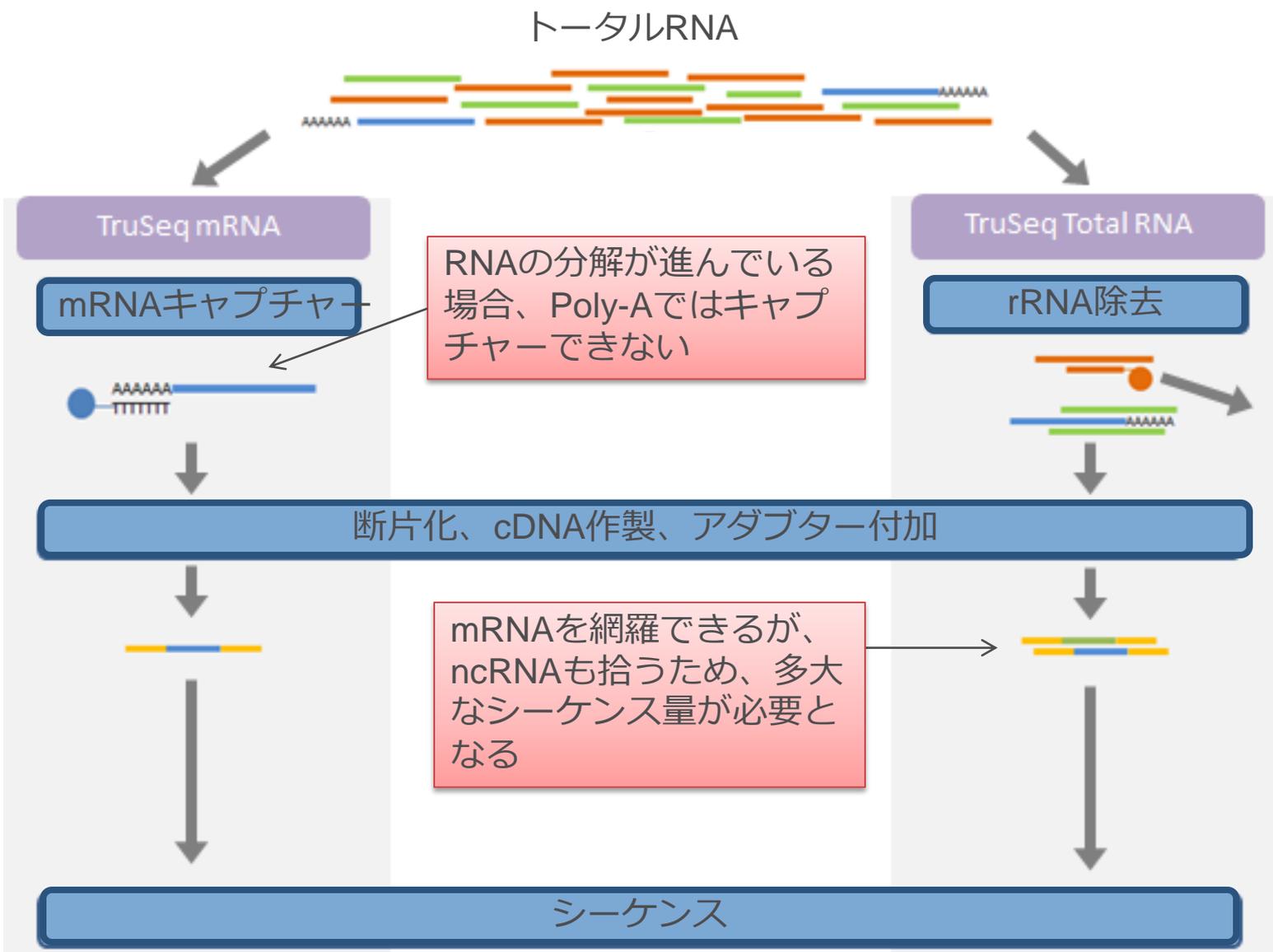
- ランダムプライマーを用い、トータルRNAから調製
- 途中でコーディング領域だけをキャプチャーすることで、mRNAに絞って解析可能

＜利点＞

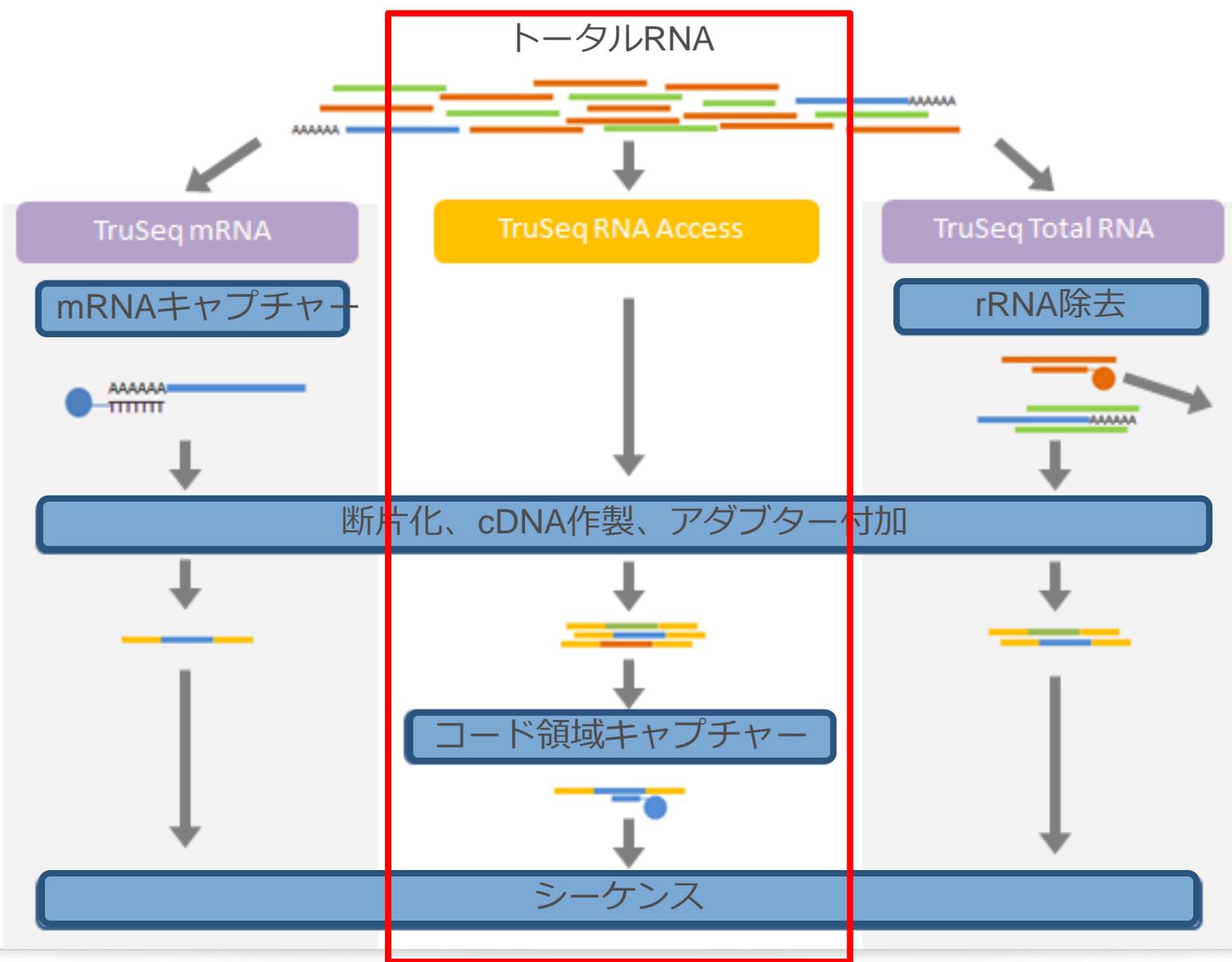
- 分解サンプルからでも、mRNAにフォーカスした解析ができる
- これまでの手法と比べて、リード量は少なくすむ
- 効率のよい解析が可能

新たな提案：TruSeq RNA Accessライブラリー調製キット

これまでの手法とその課題



新たな提案：TruSeq RNA Accessライブラリー調製キット



RNAサンプル調製キットに関するリソース

▶ サポートウェビナー

– 2014/04/18

[RNA-Seq入門（キットの選び方、実験デザイン）](#)

– 2014/03/31

[あなたのアプリケーションに適したサンプル調製キットを探す](#)

– 2015/03/20

[TruSeq DNA & RNAキットを用いたライブラリー作製におけるトラブルシューティング](#)

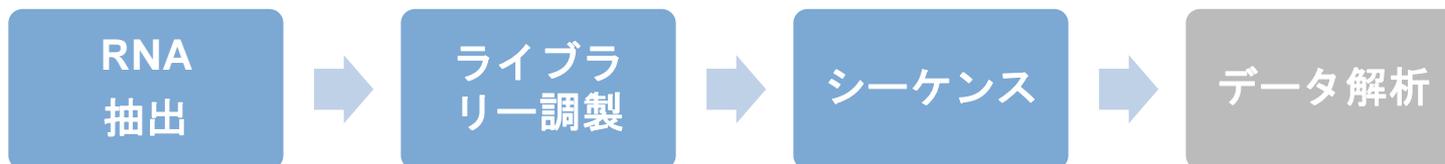
▶ イルミナホームページ:

– サンプル調整キット製品情報

- [TruSeq RNA v2](#)
- TruSeq Stranded mRNA: [LT](#), [HT](#)
- TruSeq Stranded Total RNA: [LT](#), [HT](#)
- [TruSeq RNA Access](#)



RNA-Seqの実験デザインで考慮すべき点



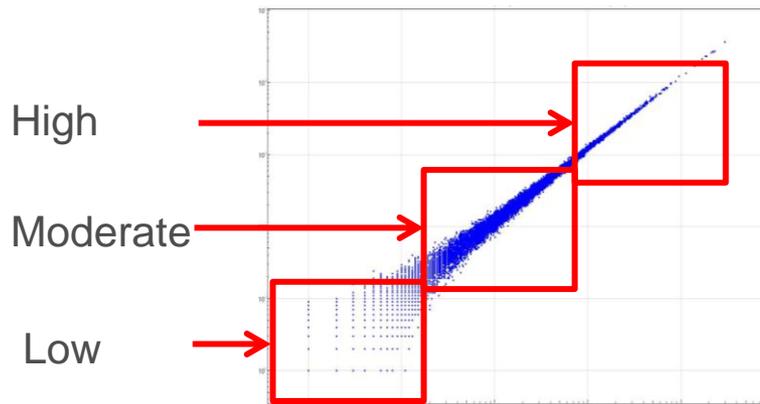
カバレッジ

カバレッジ深度 とリード数

- ▶ より高いカバレッジ = より信頼性の高い配列情報を得ることができる

```
ACGTTGACGATAGCGTCTCAGTCTGATCATAACAGTACGTTGACGATAGCGTCTCAG
CGTTGACCGTAGCGTCTCAGTCTGATCATAACAGT CGTTGACGATAGCGTCTCAG
TGACCGTAGCGTCTCAGTCTGATCATAACAGTACC TGACGATAGCGTCTCAG
TGACCGTAGCGTCTCAGTCTGATCATAACAGTACC ATAGCGTCTCAG
CGTAGCGTCTCAGTCTGATCATAACAGTACGTTGA
GCGTCTCAGTCTGATCATAACAGTACGTTGACGA
```

- ▶ RNA-Seq: シーケンスのリード数が重要!



- ▶ 遺伝子はそれぞれ異なるコピー数で存在する
- ▶ リード数が増えれば、低発現のRNAの解析も可能に

RNA-Seqのシーケンス条件

*mRNA Seq*もしくは*Total RNA Seq*の場合

ヒトRNAサンプルを用いた場合

	アプリケーション	推奨シーケンス条件 (number of reads)
発現プロファイリング解析	遺伝子発現解析	10 - 20M リード* @ 1 x 50 bp
	アレル特異的遺伝子発現解析	50 - 100M リード* @ 1 x 50 bp
トランスクリプトーム解析	スプライスバリエーション	50 - 100M リード* @ 2 x 50 ~ 2 x 100 bp
	転写産物のアセンブル	50 - 200M リード* @ 2 x 75 bp ~

* ポリAセレクションを行って作成したmRNA libraryの場合

- Total RNA library (Ribo-Zero精製) クオリティの高いRNAから作成したRNAライブラリーで2倍のリード数を、FFPEサンプルでは4倍のリード数を用いることを推奨

- あくまでも一般的な推奨条件となりますので、目的のアプリケーションに応じて、論文などをご参照いただき、シーケンス条件を設定いただくことをお勧めします
- [Encode guidelines](#)をご参照ください

RNA-Seqのシーケンス条件

Small RNAの場合

	アプリケーション	推奨シーケンス条件 (number of reads)
Small RNA	発現解析	1M リード @ 1x50 bp
	新規Small RNAの探索	10 - 20M リード @ 1x50 bp

- あくまでも一般的な推奨条件となりますので、目的のアプリケーションに応じて、論文などをご参照いただき、シーケンス条件を設定いただくことをお勧めします
- [Encode guidelines](#)をご参照ください

RNA-Seqのシーケンス条件

Small RNA

	アプリケーション	推奨シーケンス条件 (number of reads)
Small RNA	発現解析	1M リード @ 1x50 bp*
	新規Small RNAの探索	10 - 20M リード @ 1x50 bp*

1x36bp run

30bp insert

TGGAAT

Is this really an adapter?

1x50bp run

30bp insert

TGGAATTCTCGGGTGCCAAG

Yes, it is!

- アダプタートリミングを正確に行うため、インサートよりもあえて、長めにシーケンスすることを推奨

- あくまでも一般的な推奨条件となりますので、目的のアプリケーションに応じて、論文などをご参照いただき、シーケンス条件を設定いただくことをお勧めします
- [Encode guidelines](#)をご参照ください

Instrument Compatibility

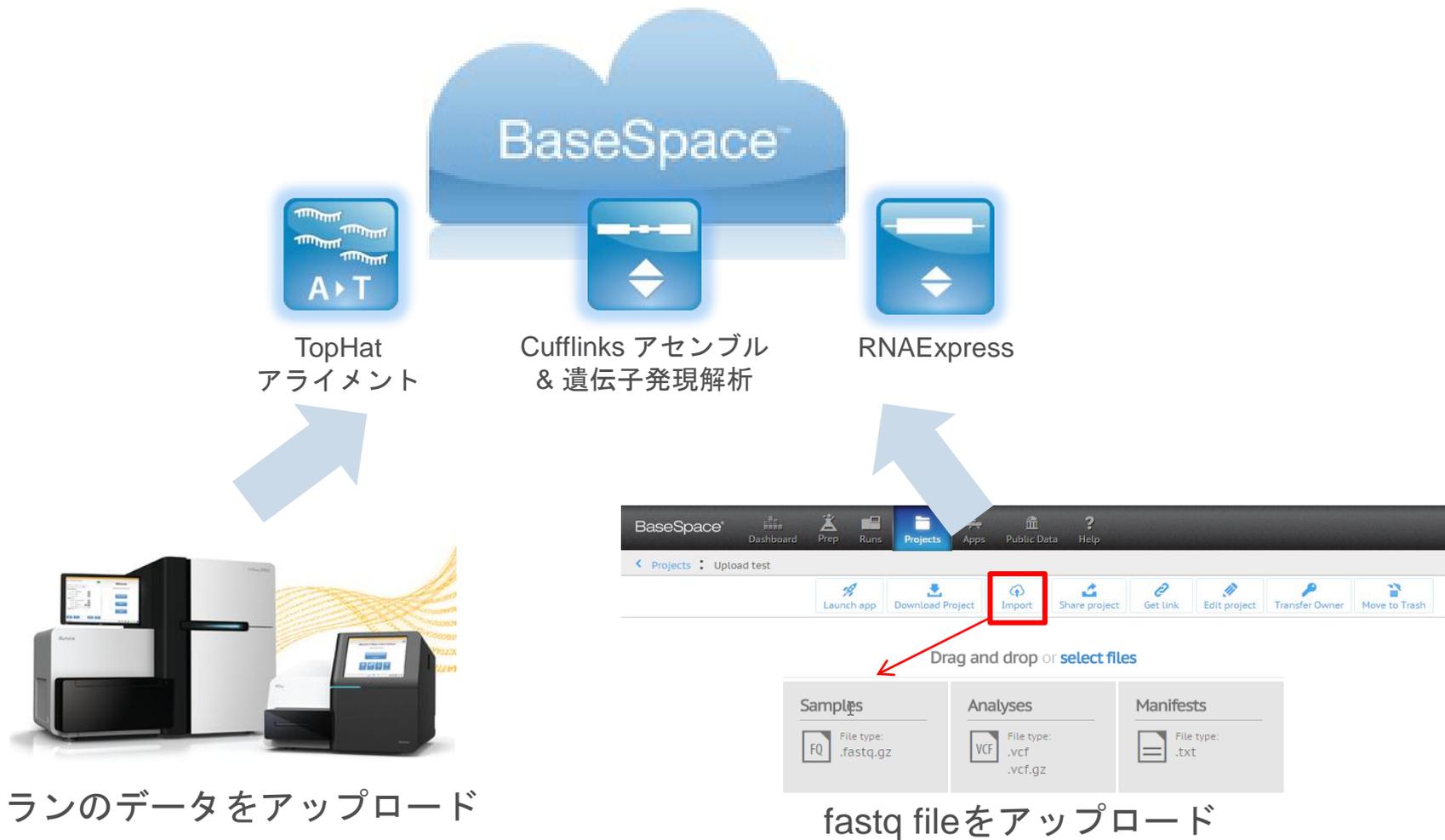
Which One to Use? How Many Samples/Run?

1度のシーケンスランで解析可能なサンプル数

	MiSeq	NextSeq 500	HiSeq 2500
発現プロファイリング解析 ¹	1 - 2	12 - 36	24 - 396
トランスクリプトーム解析 ²	-	3 - 10	8 - 96
Small RNA ³	3 - 5	25 - 80	60 - 768

^{1, 2, 3}1) 10Mリード/サンプル、2) 40Mリード/サンプル、3) 5Mリード/サンプルの場合

BaseSpaceの3つのRNA-Seq解析App



- BaseSpace上でRNA-Seqに解析に使用可能な3つのフリー解析アプリ
- 数回のクリックで、簡単に解析が行えます

BaseSpaceの3つのRNA-Seq解析App

手法	アイコン	内容
TopHat アライメント		<ul style="list-style-type: none">• 業界標準のTopHat2を使ったRNA-Seqアライメントとカウンティング• 融合遺伝子のコール（オプション）• ISSAC Varinat Callerを使用したcSNPコール• 結果はCufflinks Assembly & DE Appでさらに解析可能
Cufflinks アセンブル & 遺伝子発現解析		<ul style="list-style-type: none">• 詳細遺伝子発現差解析• 選択的転写産物のアセンブルと新規転写産物予測
RNAExpress		<ul style="list-style-type: none">• 迅速な遺伝子発現プロファイルをSTARアライメントとDESeq2で実現• 遺伝子レベルの遺伝子発現に特化

BaseSpaceの3つのRNA-Seq解析App

それぞれのアプリでできること



機能	TopHat/ Cufflinks*	RNA Express
シーケンス量のフィルター	あり	あり
シーケンスアラインメント	あり	あり
変異コール	あり	なし
融合遺伝子コール	あり	なし
転写産物アセンブル	あり	なし
遺伝子量予測	あり	なし
転写産物量予測	あり	なし
遺伝子発現差の解析	あり	あり

*[Demo data available for Tophat/Cufflinks](#): stranded mRNA & total RNA on HiSeq & NextSeq

RNA-Seq Data Analysis Software Packages

その他の解析ソフト

- **CASAVA: シングルリードのRNA-Seq解析**
 - RNA-Seqの簡単な発現プロファイリングの解析が可能
 - ノーマライズ後のリードカウントを出力
- **推奨のThird-party software***
 - [Partek Genomic Suite](#)
 - [RNA-Star](#)
 - [Tophat](#)¥[Bowtie](#)¥[Cufflinks](#)
- **BaseSpace RNA-Seq Appのご紹介**
 - [A Guided Tour](#)

*詳細はTech Noteをご参照ください: [Third-Party Analysis Software and Utilities](#)

シーケンス、データ解析に関するリソース

▶ イルミナホームページ

- [RNA Sequencing](#)
- Data Sheet: [Getting Started with RNA-Seq Data Analysis](#)
- Tech Note: [Third-Party Analysis Software and Utilities](#)
- [All Publications](#)

▶ Encode:

- [Standards, Guidelines and Best Practices for RNA-Seq: 2010/2011](#)

▶ Webiner

- 2014/10/17
- [RNA Seqをはじめよう！ BaseSpace で行う簡単NGSデータ解析](#)
- 2014/09/12
- [次世代シーケンサー（NGS）の新たなデータ解析アプローチ：BaseSpace](#)

▶ SEQanswers

ご清聴ありがとうございました!

本日セッション終了後のご質問は、
techsupport@illumina.comにお問い合わせください