27-OCT-2021

RNAシーケンスのご紹介

佐二木　健一

アプライドゲノムスペシャリスト

illumina®

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。

© 2021 Illumina, Inc. All rights reserved.

RNAは生命現象の重要な構成要素の一つ

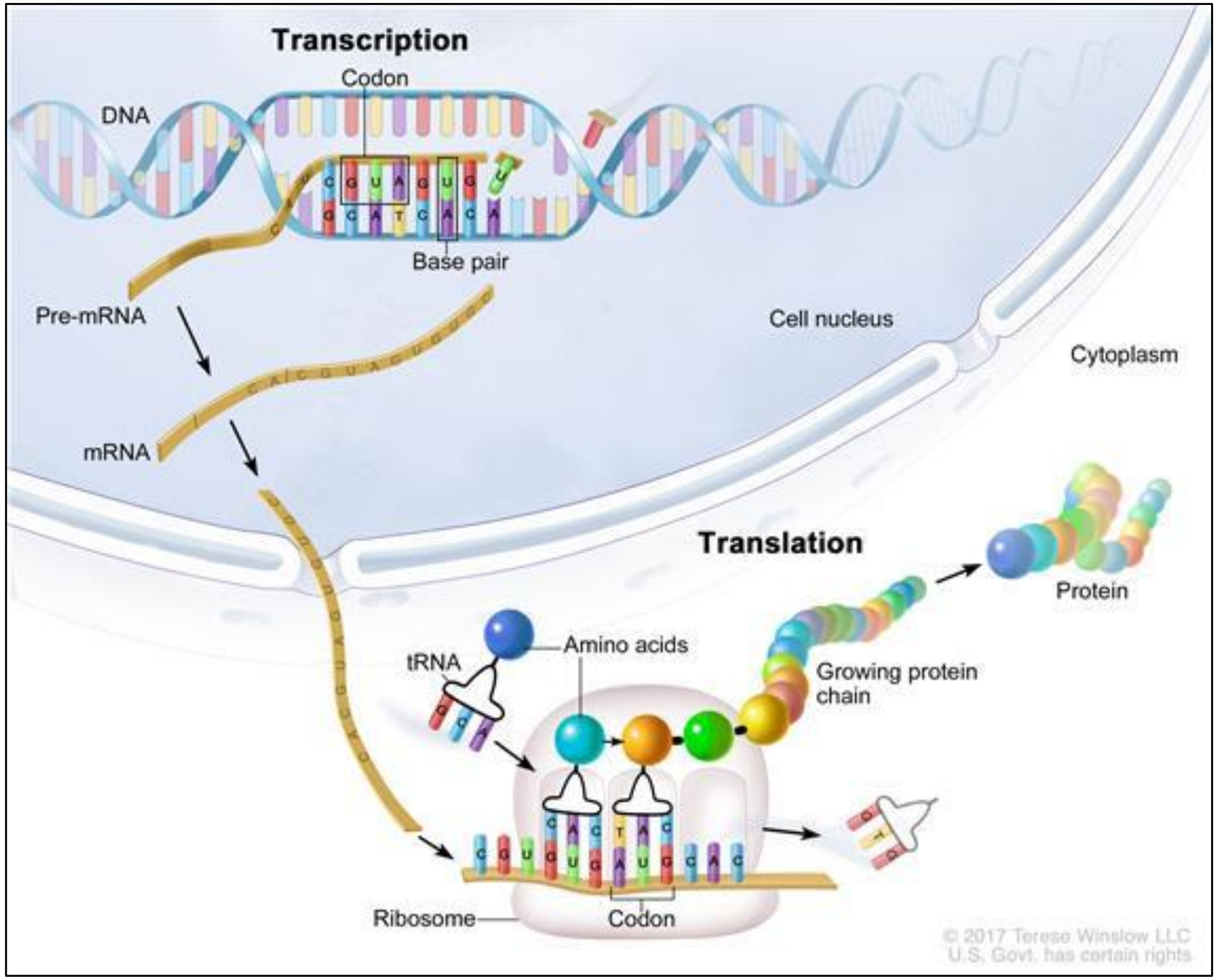


Image courteous of https://clipartxtras.com

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。

RNAは生命現象の重要な構成要素の一つ

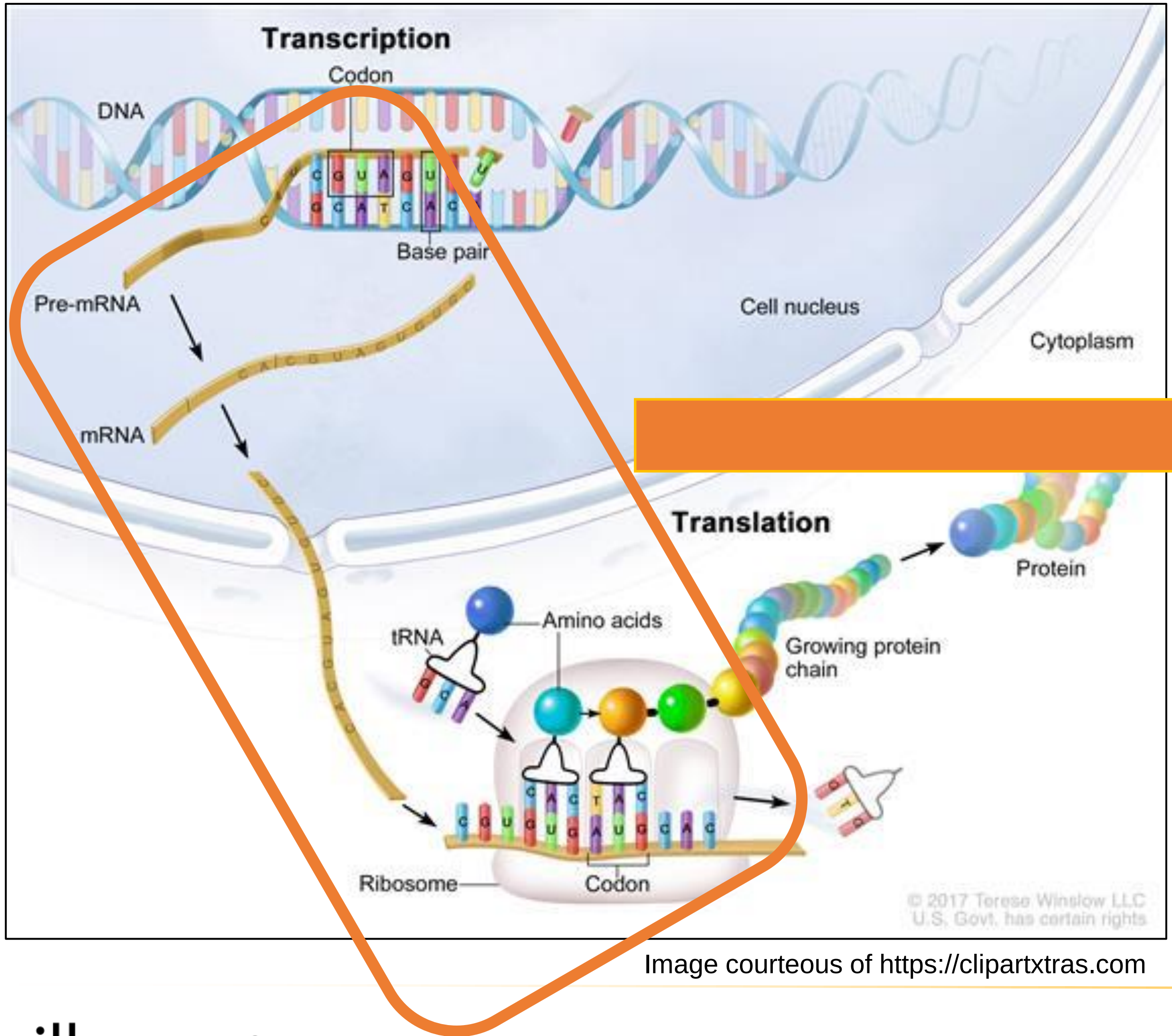


Image courteous of https://clipartxtras.com

RNA-Seq
次世代シーケンサーによる
RNAの網羅的解析

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。

RNA シーケンス
転写解析のための強力なツール

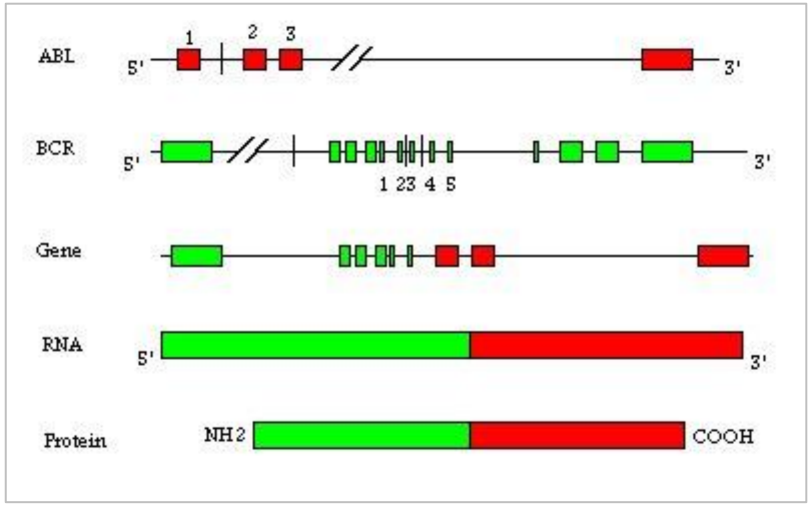
発現プロファイリング

- 単一の条件で遺伝子または転写産物の存在量を定量化する
- 異なる条件間の遺伝子または転写産物の存在量の相対的変化を測定する

転写物解析

- トランスクリプトーム全体をアセンブルします
- 既知または新規のRNA種を検出します　例 スプライスバリアント、融合遺伝子、低分子RNA、非コードRNA、ウイルスRNA

融合遺伝子（Fusion）検出



© User: A Obeidat / Wikimedia Commons / CC-BY-SA-3.0

- 融合遺伝子は、以前は独立していた2つの遺伝子の組み合わせです
- 染色体転座、欠失、または逆位が原因で発生する可能性があります
- RNA-Seqは、既知の融合と新規の融合の両方を検出するための費用効果の高い方法です。

RNAシーケンスのワークフロー

RNA抽出

目的の生物からトータルRNAを抽出します。

ライブラリー調製

シーケンス（NGS）

解析

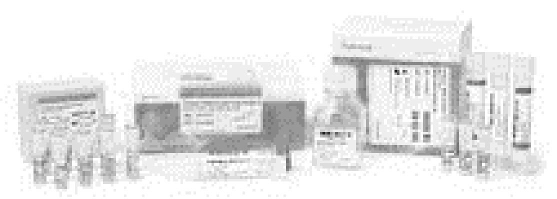
本製品の使用目的は研究に限定されます。診断での使用はできません。

RNAシーケンスのワークフロー

RNA抽出

目的の生物からトータルRNAを抽出します。

ライブラリー調製



シーケンス（NGS）

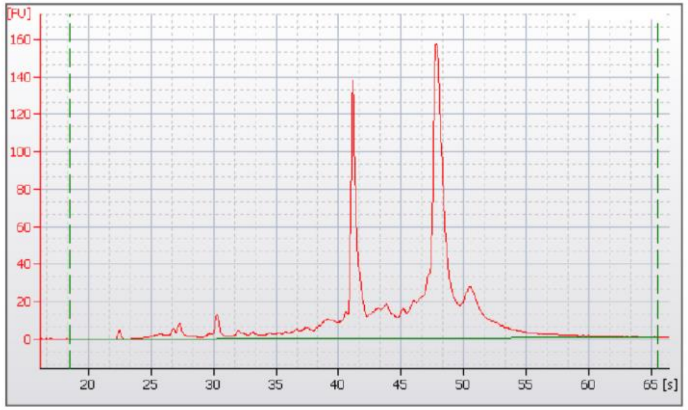
解析

RNA抽出

サンプル

- RNAはRNaseの影響を受けやすいため、収集時にサンプルを安定させ、すぐに凍結または処理することをお勧めします
- DNase処理でDNAコンタミネーションを低減してください
- DNase処理により、サンプルの純度と正確な定量が保証されます
- 特定の組織タイプまたは微生物は、溶解（Lysis）を確実に行うために別途処理が必要な場合があります

品質管理



Example of high quality UHR trace from Illumina Stranded mRNA Prep Reference Guide showing 18S and 28S rRNA peaks

- フラグメント分析法を使用した質的管理（例：バイオアナライザー）
- 蛍光光度法を使用した量的管理（例：キュービット）

RIN値 & DV_{200}

Sample	RIN	DV_{200}*
Breast Normal	2.3	77
Breast Tumor	2.7	71
Lung Normal	2.9	55
Lung Tumor	3.2	50
Colon Normal	N/A	32
Colon Tumor	N/A	39
Stomach Tumor	2.4	30
Stomach Normal	2.6	8

[Evaluating RNA Quality from FFPE Samples (illumina.com)]

- **RNA Integrity Number (RIN):**
  グラフの下の総面積に対する18Sおよび28S rRNAピークの下の面積の比率
- **DV_{200}:**
  200ヌクレオチドを超えるRNA断片の割合

RNA Input Recommendations

The protocol is optimized for 10–100 ng RNA input. It accepts 1–1000 ng purified total RNA input from high-quality RNA samples (RIN ≥ 9) and accepts 10–100 ng RNA input from low-quality RNA (RIN ≥ 2) samples or FFPE (DV200 > 55%) samples. Library performance can vary with lower input amounts and lesser quality RNA.

Example RNA Input Recommendations from Illumina Stranded Total RNA Prep Reference Guide

RNAシーケンスのワークフロー

RNA抽出

目的の生物からトータルRNAを抽出します。

ライブラリー調製

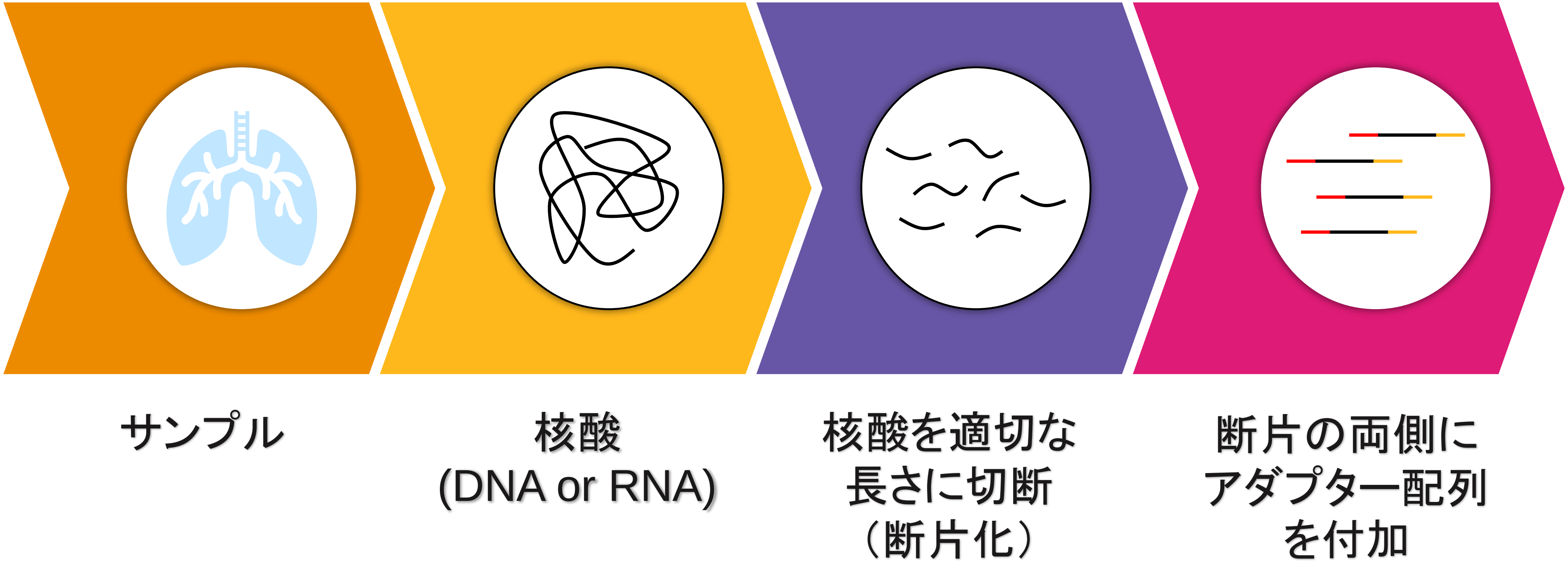
シーケンス（NGS）

解析

illumina®

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。

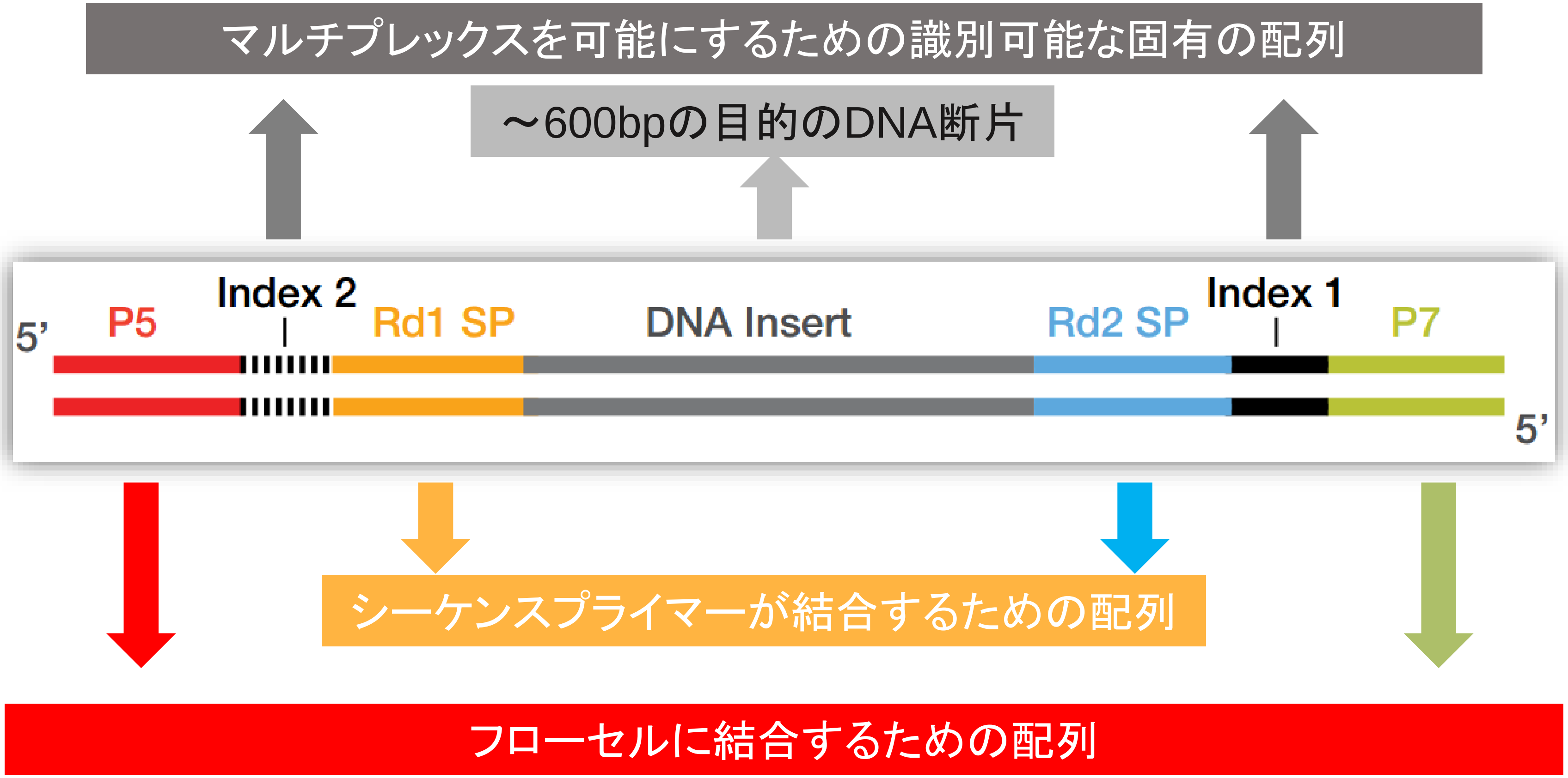
そもそもライブラリーとは？
サンプルの前処理



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。

ライブラリーの構造と役割

全てのイルミナシーケンサーで共通

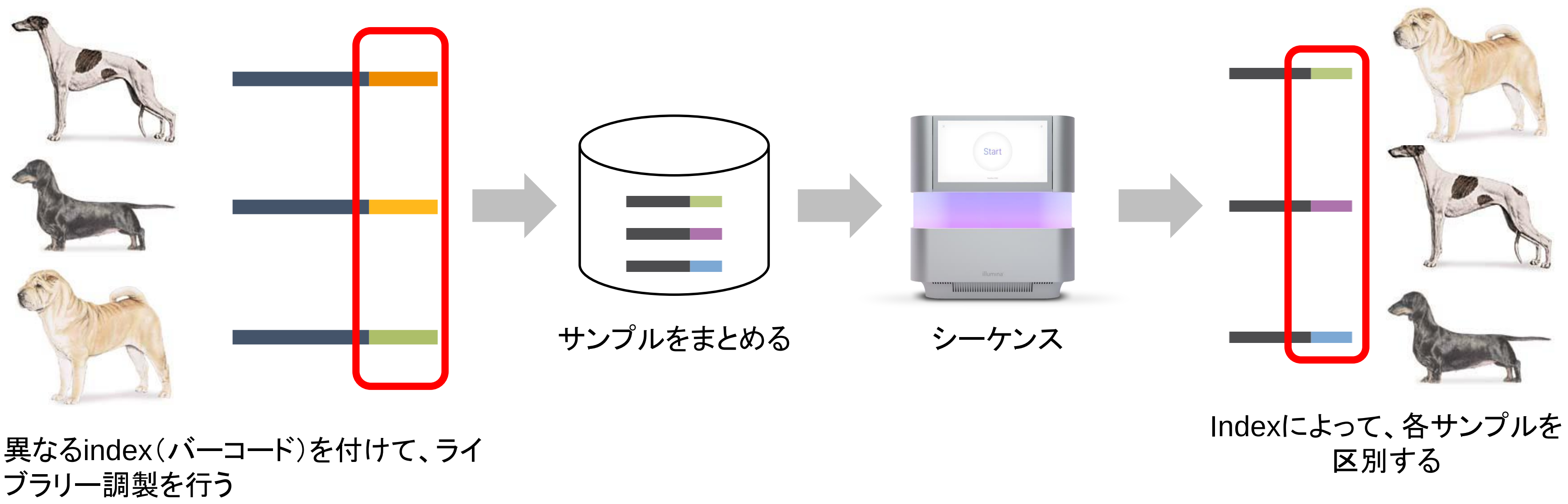


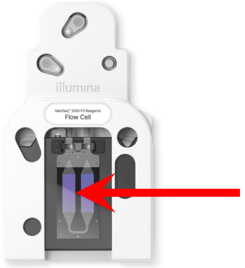
illumina®

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。

マルチプレックス法

- 異なる配列のインデックスを利用することで、1回のランで多サンプルを同時にシーケンスできる

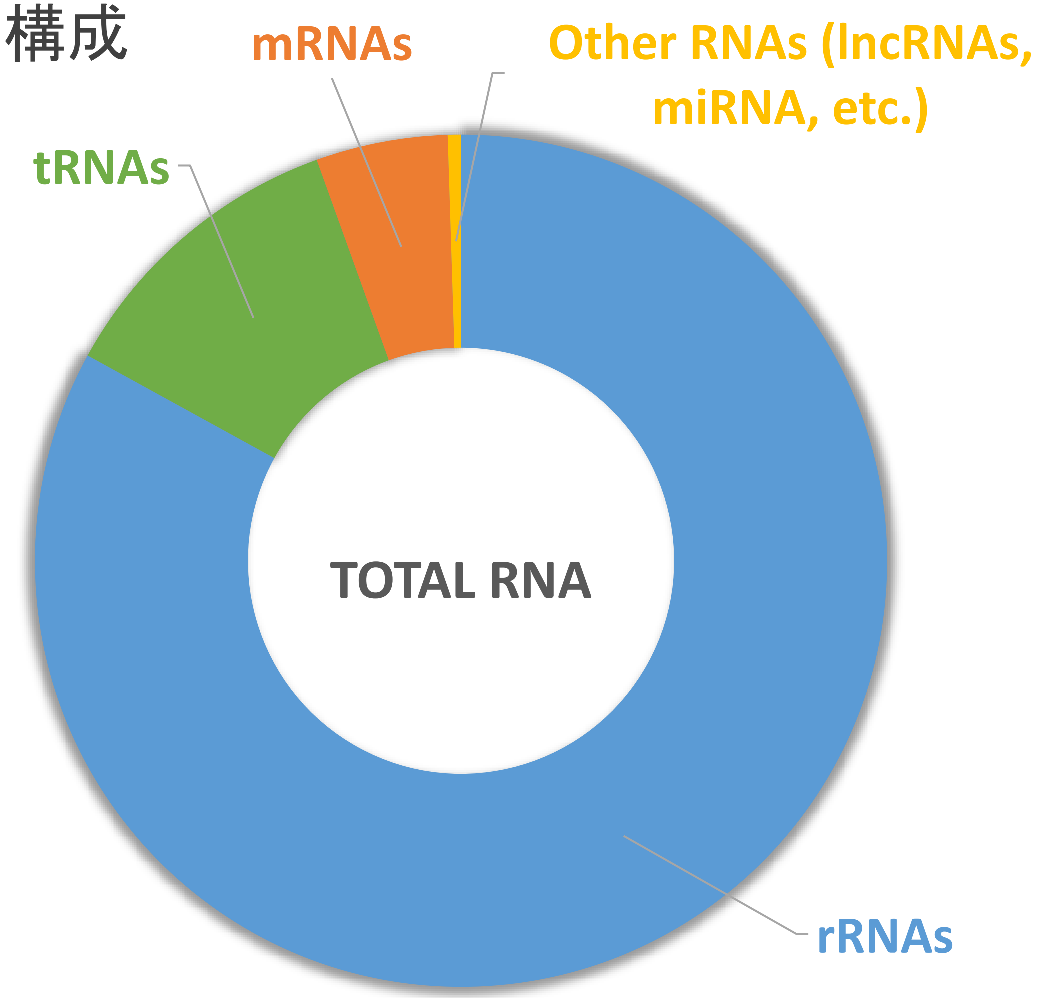




1レーンに複数サンプル混ぜることができる

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。

RNAライブラリー
哺乳類のRNAの構成



- 大量に存在するrRNAとtRNAは一般的にはRNAシーケンシングから除外されます
- mRNA とsmall RNAを一般的にはターゲットにして解析を行います

Non-coding RNA: what is functional and what is junk? (2015) AF Palazzo and ES Leeより改変

ターゲットとするRNA種を得るために

1

Ribosomal RNA の除去

- rRNAを標的とするプローブを使って、サンプルから除去します
- コーディング領域と非コーディング領域の検出

2

Poly A の選択

- オリゴd(T)を使って、ポリAテールを持つmRNAを選択します
- ポリAテールを持つ転写産物の検出

3

cDNA 合成 & ターゲット濃縮

- RNAからcDNAへの変換
- 特定のプローブを使用して、ターゲット領域を選択します
- エクソームパネルなど

RNA シーケンシングソリューション

1

2

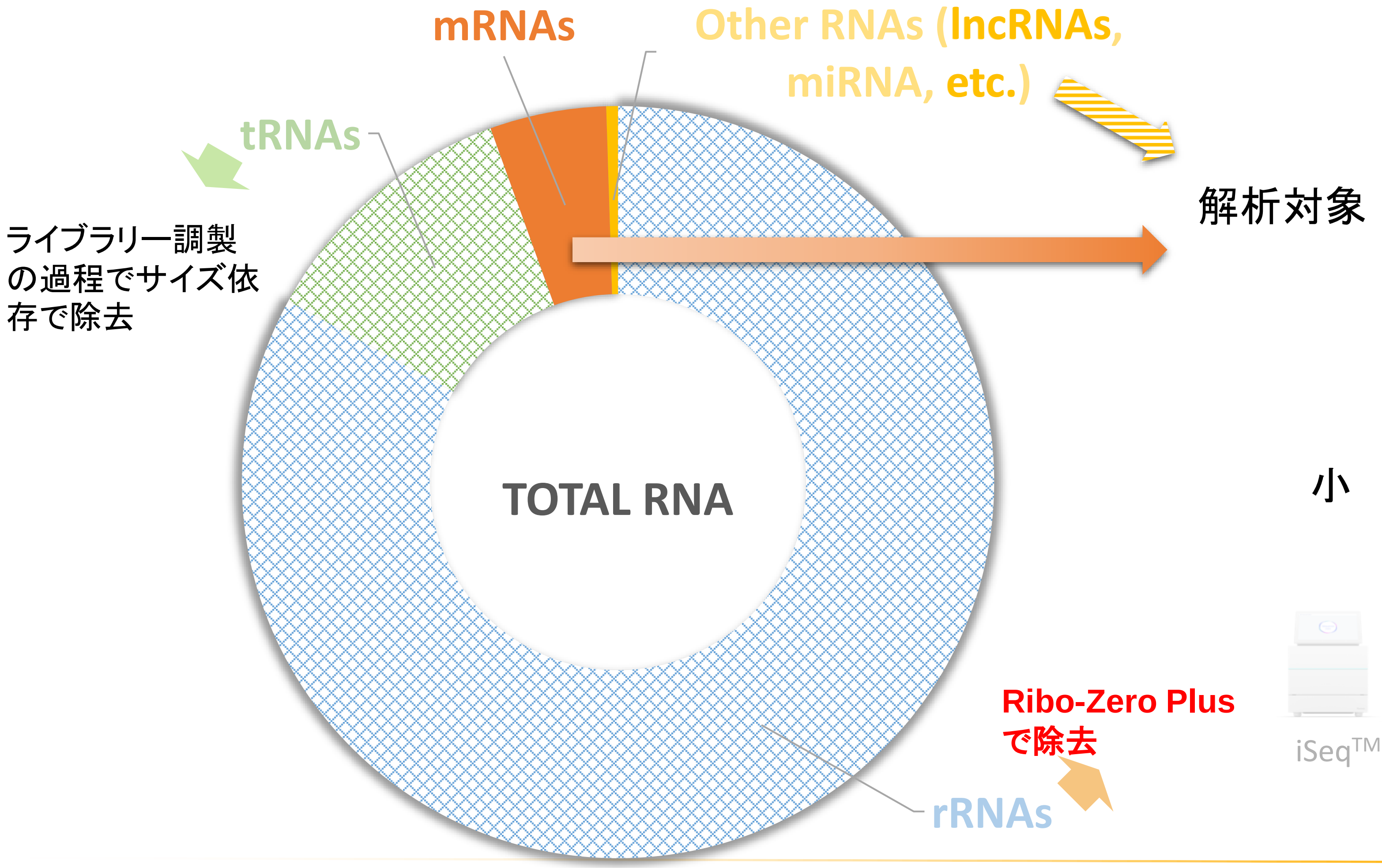
3

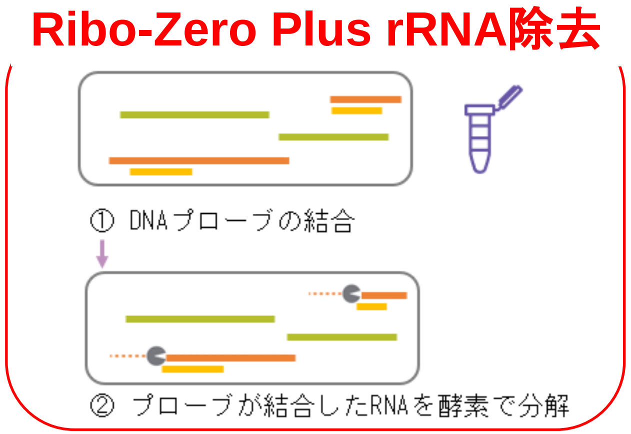
	Ribosomal RNA の除去	Poly A の選択	cDNA 合成 & ターゲット濃縮
	トータル RNA	**mRNA**	**ターゲット濃縮**
	最大の探索力	**コスト効率の高い発現解析**	**ターゲットを絞った検出と発見**
検出対象	コーディング領域と非コーディング領域	ポリAテールを持つ転写産物	ターゲット領域
劣化したサンプル/ FFPE	✓		✓
適したアプリケーション	• グローバルな発現プロファイリング • ncRNAの発見と特性評価 • 選択的スプライシングの検出 • バイオマーカーの同定	• 差次的発現プロファイリング • アイソフォームの識別 • 対立遺伝子特異的発現 • 低発現mRNAの検出	• 対立遺伝子特異的発現 • 劣化したサンプル • バイオマーカースクリーニング
Illumina RNA調製キット (16 & 96 サンプルキット)	**Illumina Stranded Total RNA Prep** with Ribo-Zero Plus	**Illumina Stranded mRNA Prep**	**Illumina RNA Prep with Enrichment**

illumina®

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。

1. キット名：Illumina Stranded Total RNA Prep with Ribo-Zero™ Plus





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。

Illumina Stranded Total RNA Prep ワークフロー

高いライゲーション反応を活用して、高品質でシンプルな調製を提供します

rRNA 除去

過多な転写産物の除去により、目的の標的転写産物に焦点を合わせます

断片化

ILMNシーケンスに最適な長さに断片化されます

cDNA 合成

逆転写によりRNAをcDNAに変換します
cDNA試薬はキットに含まれます

ライゲーション[1] & 増幅

ライゲーションによりインデックスが付加され、出来たライブラリーをPCRで増幅します

シーケンス

ILMNシステムと完全に適合するように設計されています

1. Quant, qual and normalization steps not shown

Illumina Stranded Total RNA Prepのワークフロー

Illumina Stranded Total RNA Prep

LIGATION WITH RIBO-ZERO PLUS

Total RNA	
リボソームRNAの除去	
RNA断片化と変性	
First Strand cDNA合成	所要時間 7時間[1]
Second Strand cDNA合成	
3'末端のアデニル化	操作時間 3時間未満[1]
アンカーライゲーション	
断片のクリーンアップ	
ライブラリー増幅	
ライブラリーのクリーンアップ	
定量とノーマライゼーション	
シーケンス	

1. 最大24サンプルを手作業で処理した際の時間で、定量とノーマライゼーションは含まれない。

Ribo-Zero Plus：酵素を用いたrRNA除去法

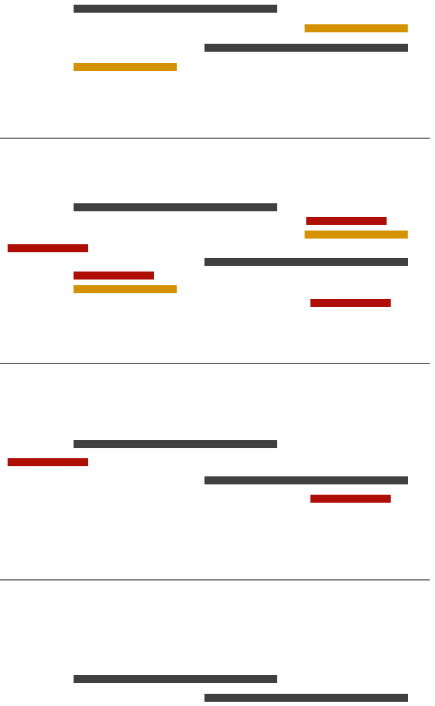
ロバストな手法により、多様なサンプルを扱うラボの柔軟性が向上

Total RNA

1 プローブのハイブリダイゼーション

2 rRNA除去

3 プローブの除去



興味のあるRNA　Ribo-Zeroプローブ　大量のrRNA

Ribo-Zero Plusにより単一チューブで複数生物種に対応

多様なサンプルを扱うラボでの研究にシームレスな柔軟性

- 細胞質およびミトコンドリアrRNA：ヒト、マウス、ラット
- グロビン転写産物
- バクテリアrRNA（グラム陽性菌/陰性菌）

除去ターゲット	ターゲットrRNA
ヒトの細胞質rRNA	28S、18S、5.8S、5S
ヒトのミトコンドリアrRNA	12S、16S
ヒトのβグロビン転写産物	HBA1、HBA2、HBB、HBG1、HBG2
マウスおよびラットのrRNA	16S、28S
グラム陰性菌rRNA	大腸菌：5S、16S、23S
グラム陽性菌rRNA	枯草菌：5S、16S、23S

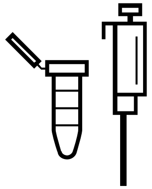
illumina®

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。

Illumina Stranded Total RNA Prepのデータ

特に低いインプット量でのワークフローを強化

SAMPLE PROFILE



Universal Human Reference

SAMPLE INPUT AMOUNT

Total RNAで1ng、10ng、100ng

SEQUENCER

NextSeq™ 550システム

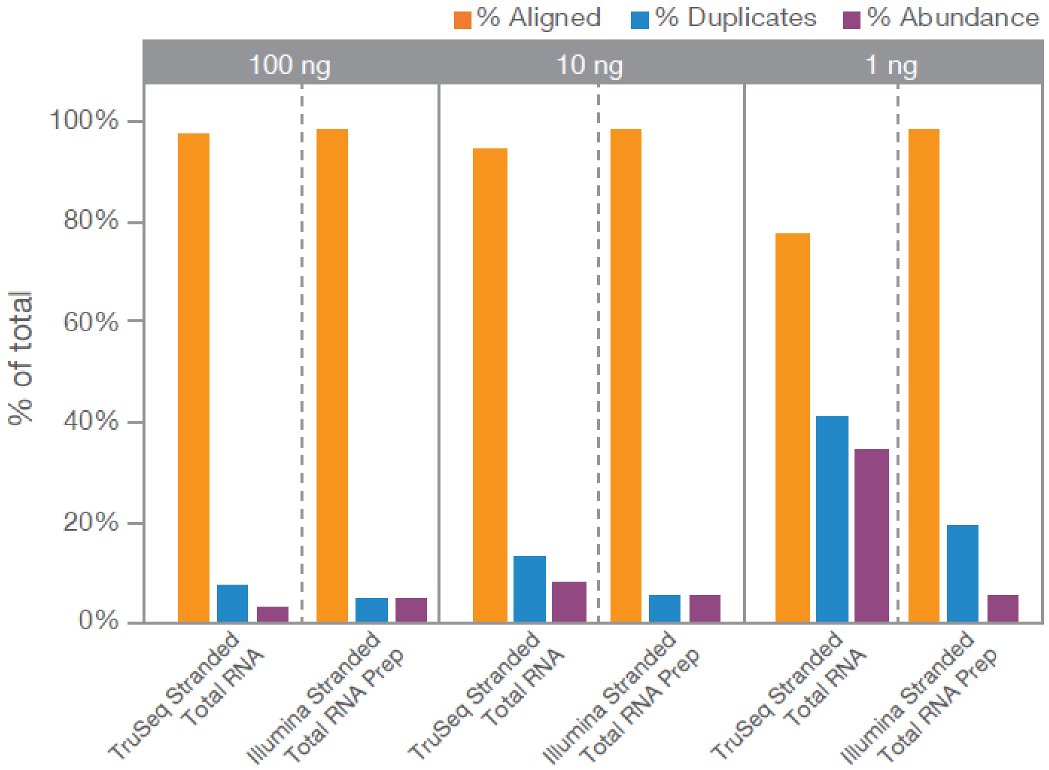
Reads

30Mリードにサブサンプリング*

解析

BaseSpace RNA-Seq Alignment v 2.0.1アプリ

1ngおよび100ngのRNAインプット量で**優れた性能**を発揮
（従来のTruSeqとの比較）

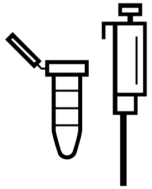


*4Mリードにサブサンプリングしてduplicateの割合を算出

Illumina Stranded Total RNA Prepのデータ

FFPEサンプルから均一で高いトランスクリプトームカバレッジ

SAMPLE PROFILE



FFPE

DV_{200} = 55%

SAMPLE INPUT AMOUNT

Total RNAで10ng
および100ng

SEQUENCER

NovaSeq™ 6000システム

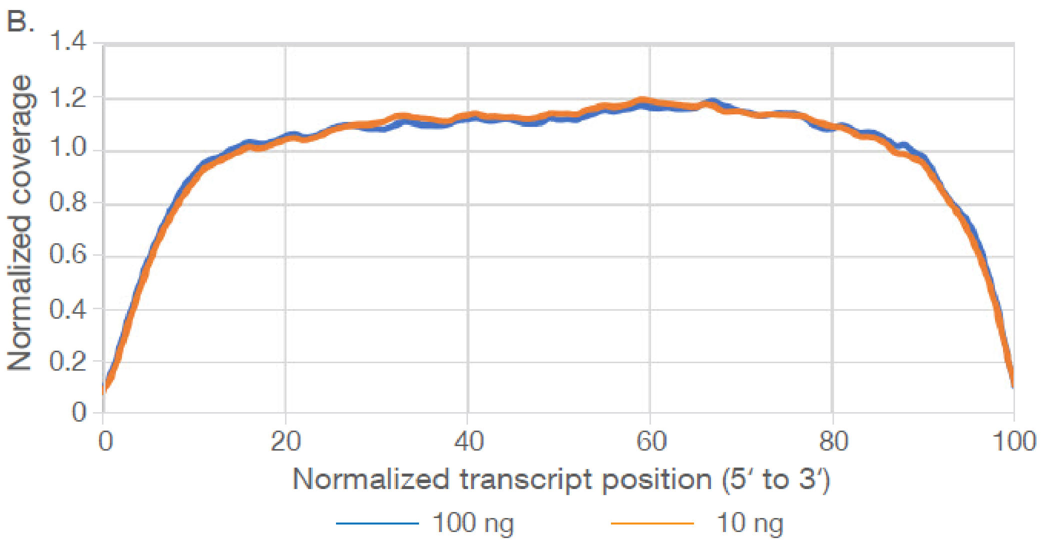
Reads

50Mリード

解析

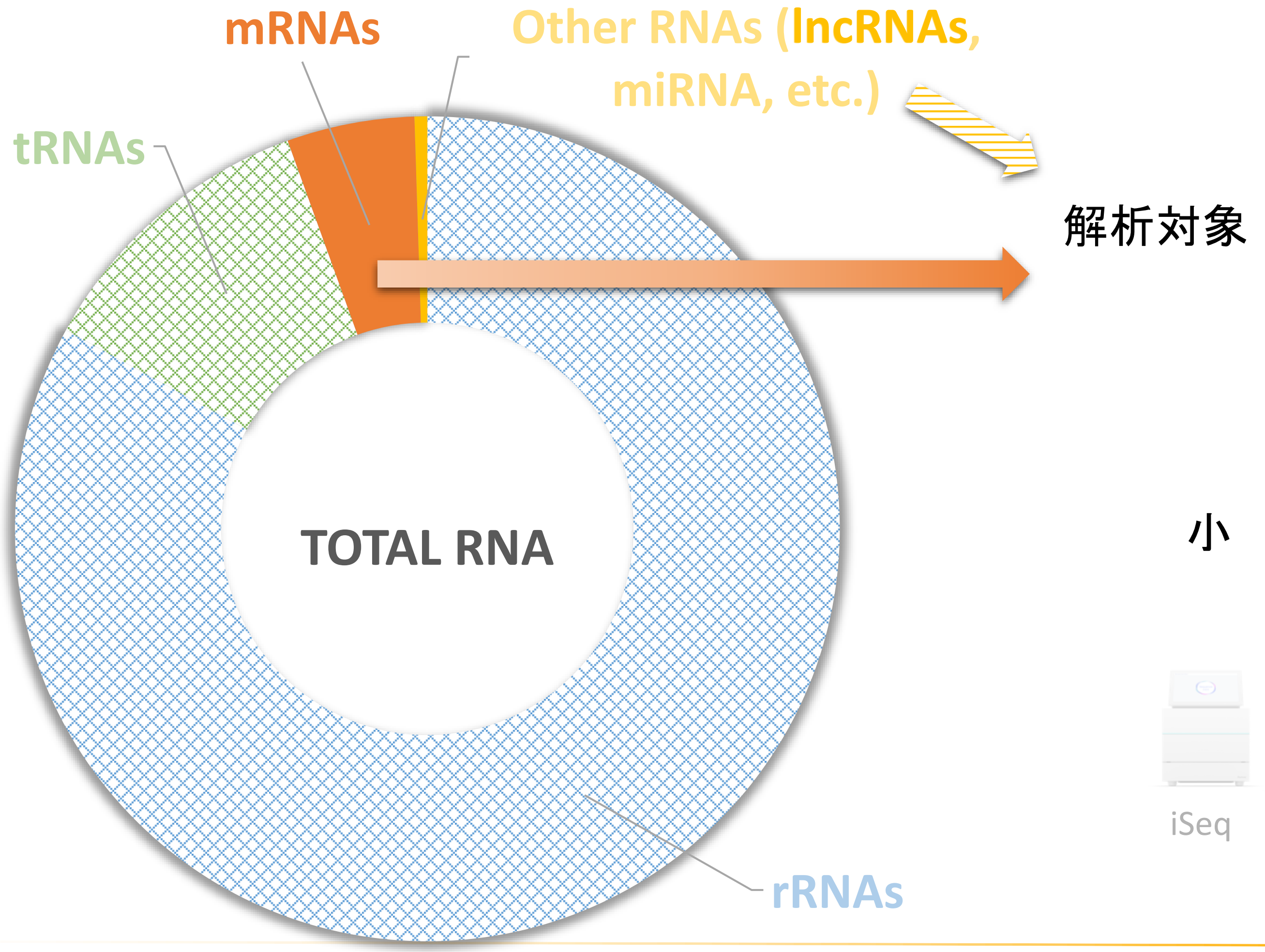
BaseSpace RNA-Seq
Alignment v 2.0.1アプリ

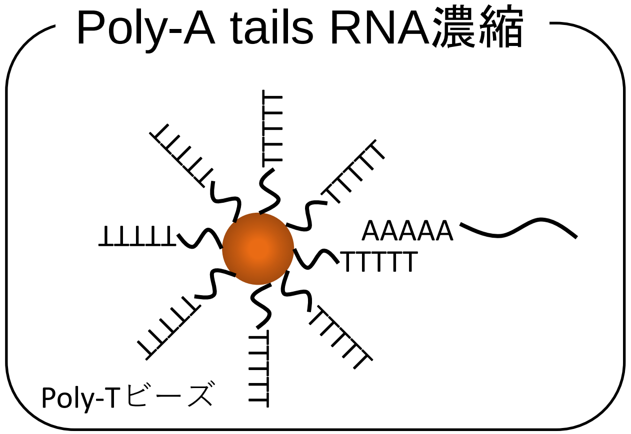
10ngおよび**100ngのRNA**インプット量で、**FFPE** RNAで
高いカバレッジ均一性



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。

2. キット名：Illumina Stranded mRNA Library Prep





illumina®

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。

Illumina Stranded mRNA Prep ワークフロー

高いライゲーション反応を活用して、高品質でシンプルな調製を提供します

PolyA選択	断片化	cDNA 合成	ライゲーション[1] & 増幅	シーケンス

mRNAの濃縮により、目的の標的転写産物に焦点を合わせます	ILMNシーケンスに最適な長さに断片化されます	逆転写によりRNAをcDNAに変換します cDNA試薬はキットに含まれます	ライゲーションによりインデックスが付加され、出来たライブラリーをPCRで増幅します	ILMNシステムと完全に適合するように設計されています

1. Quant, qual and normalization steps not shown

Illumina Stranded mRNA Prepのワークフロー

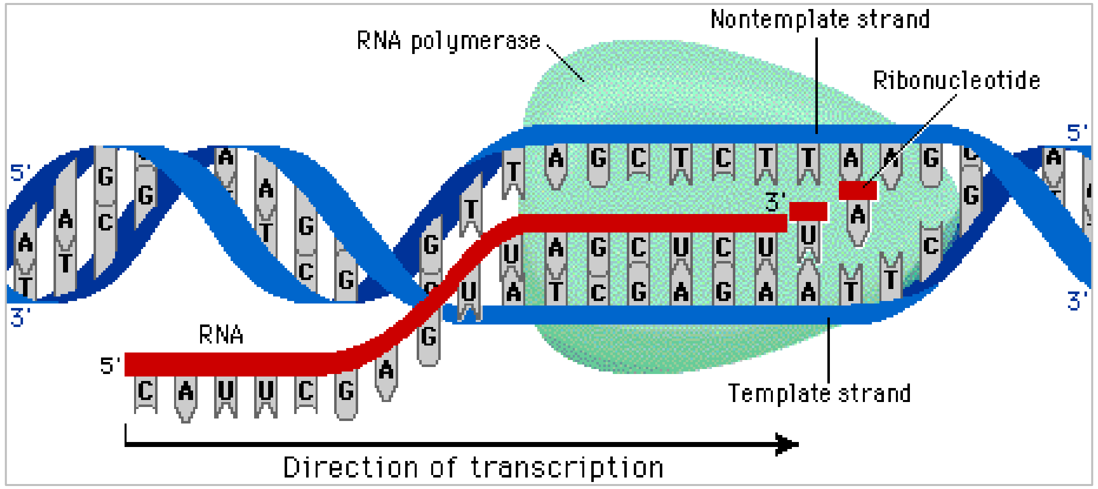
Illumina Stranded mRNA Prep

LIGATION

Total RNA	
mRNA精製と断片化	所要時間 6.5時間[1] 操作時間 3時間未満[1]
First Strand cDNA合成	
Second Strand cDNA合成	
3'末端のアデニル化	
アンカーライゲーション	
断片のクリーンアップ	
ライブラリー増幅	
ライブラリーのクリーンアップ	
定量とノーマライゼーション	
シーケンス	

1. 最大24サンプルを手作業で処理した際の時間で、定量とノーマライゼーションは含まれない。

ストランド情報/Strandednessとは?



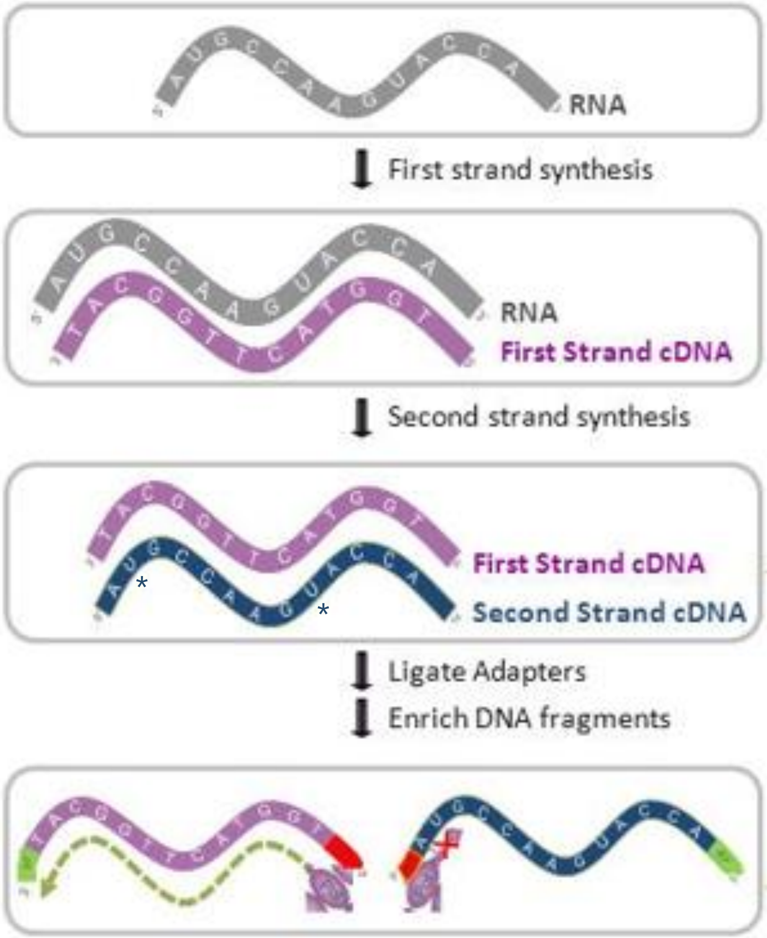
- ストランドネスは、転写産物がどのDNAストランドに由来するかを決定します
- アンチセンス発現の検出を可能にして、トランスクリプトアノテーションの信頼性を高めます
- 次のアプリケーションにとってストランド情報は重要です：トランスクリプトベースのアセンブリ、完全なアノテーション、スプライスバリエーション研究

cDNAのセカンドストランド合成ステップで導入されるストランド情報

1 Purify and Fragment Messenger RNA
Hands-on: 46 minutes
Total: 1 hour 21 minutes
Reagents: BBB, BWB, ELB, EPH3, RPBX

2 Synthesize First Strand cDNA
Hands-on: 5 minutes
Total: 50 minutes
Reagents: FSA, RVT

3 Synthesize Second Strand cDNA
Hands-on: 35 minutes
Total: 1 hour 40 minutes
Reagents: 80% EtOH, AMPure XP, RSB, SMM



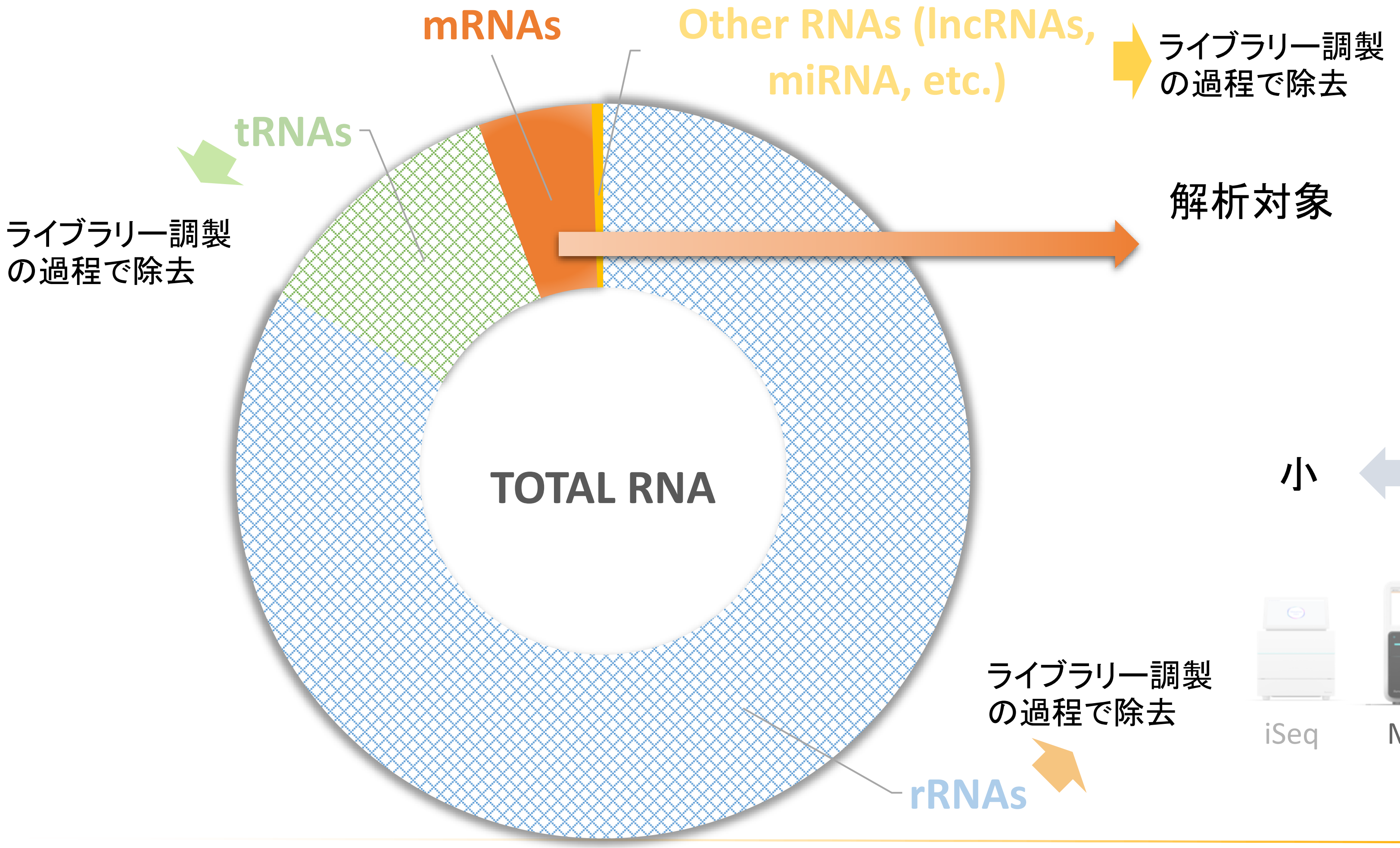
- 二本鎖を合成して平滑末端cDNAを生成します
- dUTPはdTTPの代わりにセカンドストランドに組み込まれています

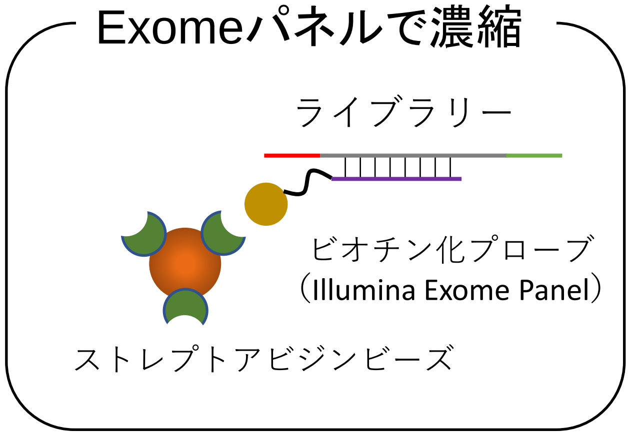
- ポリメラーゼはdUTPを組み込むことができません
- セカンドストランドは増幅ステップ中に抑制されます

illumina

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。

3. キット名：Illumina RNA Prep with Enrichment
Exome Panelを使った場合





illumina

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。

Illumina RNA Prep with Enrichment ワークフロー

超高速RNA濃縮のためにタグメンテーションケミストリーを活用[1]

cDNA 合成 | タグメンテーション | ターゲット濃縮 | 増幅 | シーケンス

cDNA 合成: 逆転写によりRNAをcDNAに変換します
cDNA試薬はキットに含まれます

タグメンテーション: シングルステップの断片化とインデックスの付加

ターゲット濃縮: ハイブリッドキャプチャベースのターゲット濃縮

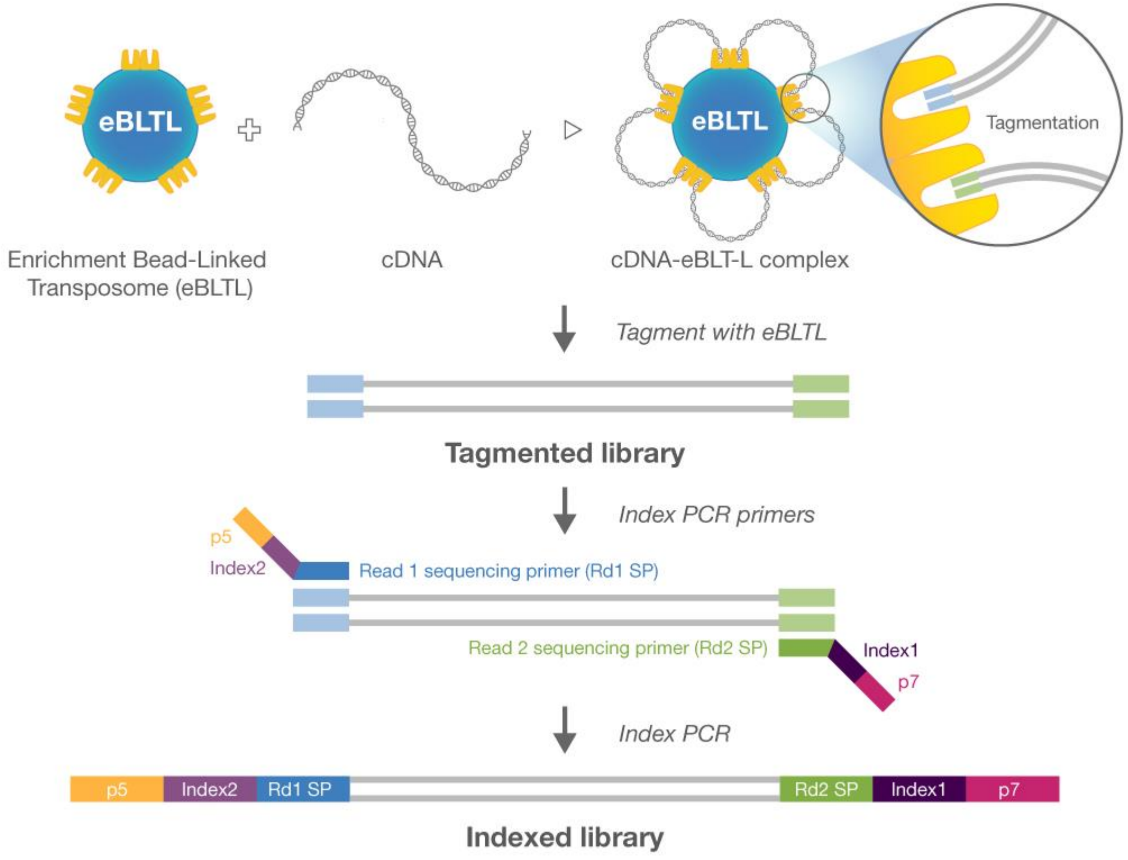
増幅: 出来たライブラリーをPCRで増幅します

シーケンス: ILMNシステムと完全に適合するように設計されています

1. Quant, qual and normalization steps not shown

タグメンテーション

RNA濃縮用に拡張されたビーズ結合トランスポゾン（BLT）

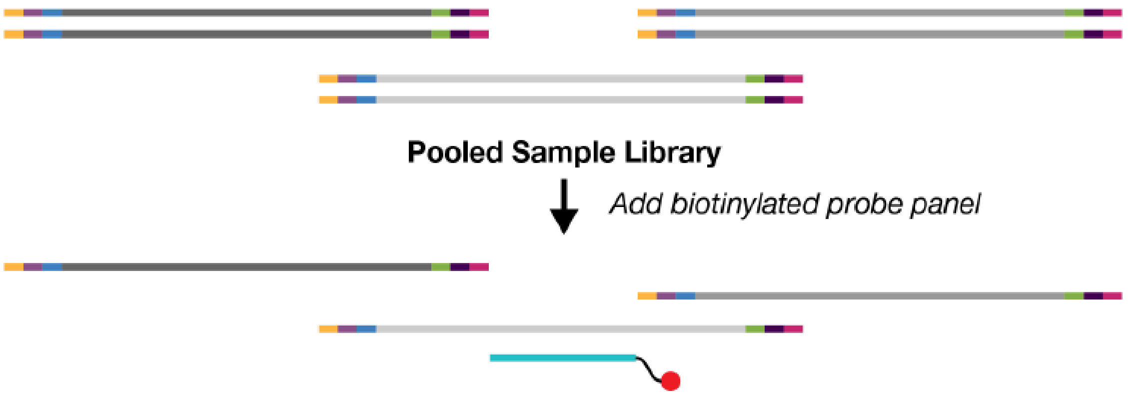


本製品の使用目的は研究に限定されます。診断での使用はできません。

ターゲット濃縮

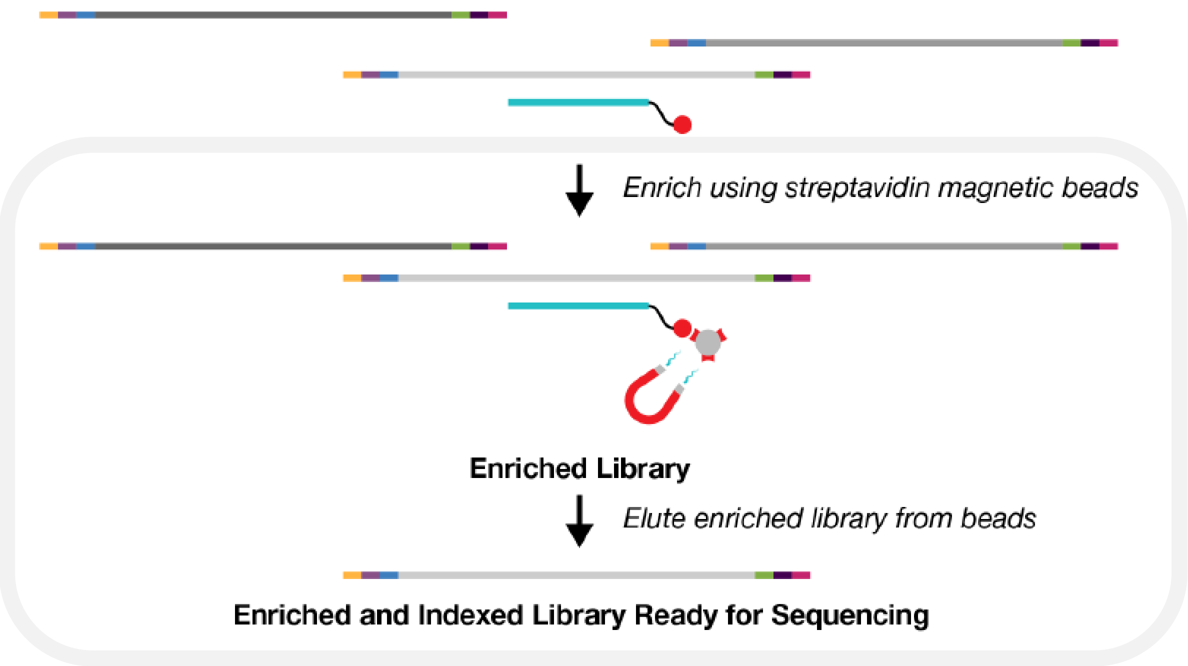
プローブをハイブリダイズ

プールされたライブラリにキャプチャプローブを追加して、関心領域を標的とします



ライブラリーの濃縮

標的にしたライブラリー断片にハイブリダイズしたマグネットビーズ濃縮プローブ



RNA 濃縮用パネル

Illumina RNA Prep with Enrichment は様々なパネルに対応します

イルミナのテスト済みパネル	現在利用可能なその他のオプション
Illumina Exome Covers >98% RefSeq Exome including >21k target genes, >214k target exonic regions	**Illumina Concierge Custom Panels** Design support available for human and non-human custom RNA enrichment panels NOTE: no performance guarantee on custom panels
Illumina Respiratory Virus Panel v2 7,800 probes to detect common respiratory viruses, recent flu strains, and SARS-CoV-2	**Use with 3rd Party Panels** No known chemistry limitations No Illumina testing or performance guarantee Modifications to protocol may be required
Respiratory Pathogen ID/AMR Panel Detects >280 pathogens contributing to pneumonia >1200 Antimicrobial Resistance markers	

RNAシーケンスのワークフロー

RNA抽出

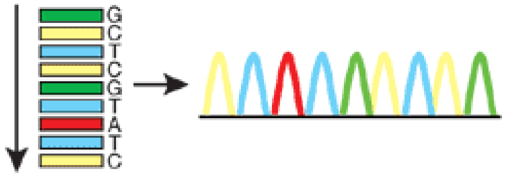
目的の生物からトータルRNAを抽出します。

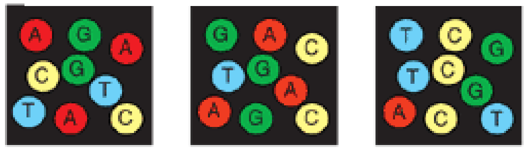
ライブラリー調製

シーケンス（NGS）

解析

リード長、リード数とデータ量の関係





	サンガーシーケンサー（1リード/キャピラリー）	次世代シーケンサー（1リード/クラスター）
リード長	1,000 bp x1	150 bp x2
X	X	X
リード数	96	10,000,000,000
II	II	II
データ量	96 Kb	3 Tb

1回で多数のリードを解読

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。

RNA-Seq解析でイルミナシーケンサーを選ぶメリット

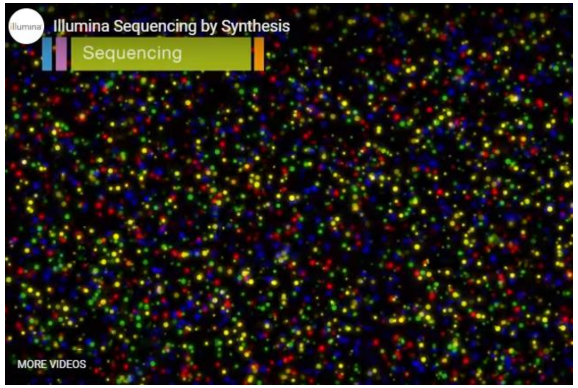
一度に解析できるリードの数が多い

↓

解析できるライブラリー分子数が多い

高感度／多検体処理

*イメージ図



例）NextSeq1000/2000
フローセル

約12億クラスター

Check!

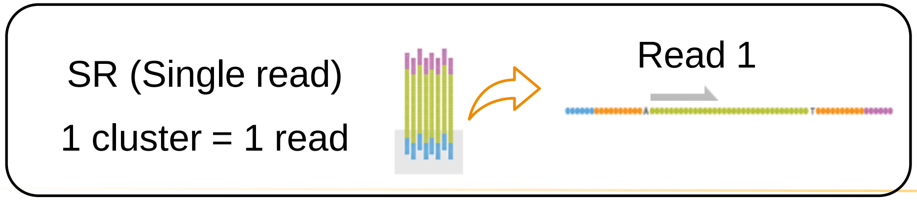
目的のアプリケーションによって推奨必要リード数が異なる

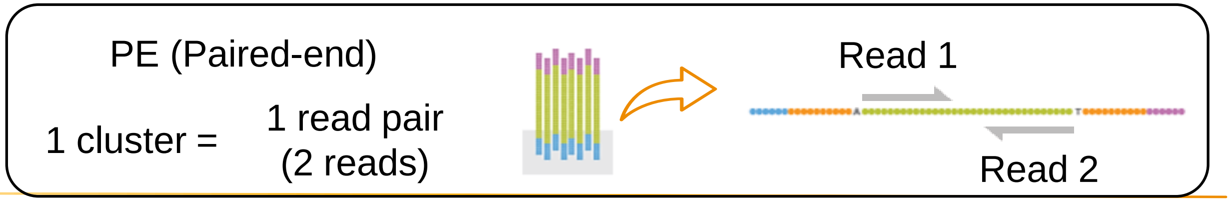
例) Guideline from ENCODE2 RNA-seq experiments

安 ← サンプルあたりのコスト → 高

アプリケーション	遺伝子発現プロファイリング解析		遺伝子発現プロファイリング解析 + 新規転写産物探索	
目的例	• 簡単な発現プロファイリング	• より広い遺伝子発現プロファイリング • アイソフォーム毎の解析 • 融合遺伝子検出	• *De novo* assembly • 詳細な遺伝子発現プロファイリング	• FFPE由来のサンプルからの*De novo* assembly
シーケンス条件	SR 35 - 50 bp	PE 50 - 75 bp	PE 75 - 150 bp	PE 75 - 150 bp

クラスター数　1000万　3000万　5000万　1億　2億





illumina

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。

現在のイルミナ次世代シーケンサーラインナップ
2021年10月現在、フローセルごとの値

2020年11月発売　2020年1月発売

システム	iSeq™ 100	MiniSeq™	MiSeq™	NextSeq™ 550	NextSeq™ 1000	NextSeq™ 2000	NovaSeq™ 6000
データ量（最大値）	1.2 Gb	7.5 Gb	15 Gb	120 Gb	120 Gb	360 Gb	3 Tb
クラスター数（最大値）	400万	2500万	2500万	4億	4億	12億	100億
ラン時間（時間）	19	24	56	29	29	48	44
発売年	2018	2016	2011	2014	2020	2020	2017

illumina

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。

新製品、NextSeq 1000/2000

NextSeq 2000

4,422万円

NextSeq 1000

2,772万円

AUG
27

最新機器 NextSeq 1000 のご紹介と MiSeq からのアップグレードプログラム

2020/08/27

イルミナ株式会社　営業本部 技術営業部　鈴木 健介

ウェビナー録画はこちら

P3 Flow Cell

12億クラスター 360Gb

最大データ量

P2 Flow Cell

4億クラスター 120Gb

最大データ量

P1 Flow Cell

1億クラスター 30Gb

最大データ量

P3はNextSeq2000のみで使用可能

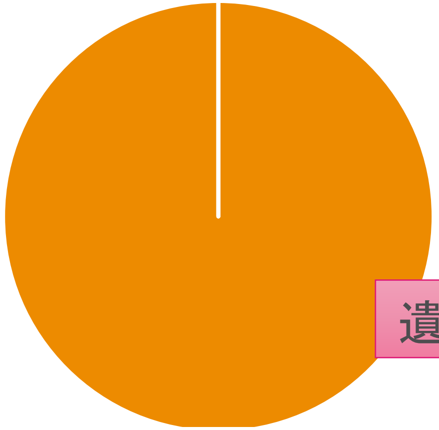
illumina

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。

シーケンス試薬利用例

フローセルのデータ量を最大限に利用するマルチプレックス解析

例）NextSeq 1000/2000　P2



~4億クラスター/120 Gb

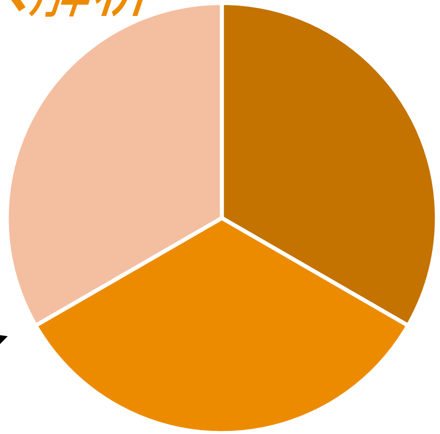
１枚のフローセルで期待されるクラスター数

１枚のフローセルで期待される総データ量
[クラスター数 x リード長]

新規転写産物探索

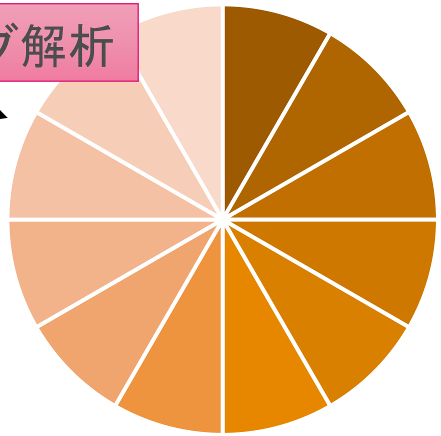
遺伝子発現プロファイリング解析

3-plex



サンプルあたり
~1.3億 clusters/40 Gb

12-plex



サンプルあたり
~0.33億 clusters/10 Gb

illumina®

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。

RNA-Seq NextSeq 2000

アプリケーション: スプライスバリアントの解析

サンプルごとのリード数: 1億リード

リード長: 2 x 50 bp

装置: NextSeq 2000

シーケンシング試薬: P2試薬 100サイクル

全データ量:

2 x 50 bp x 4億リード = 40ギガベース

1ラン当たりのサンプル数: 8サンプル

	NextSeq 2000 P2 試薬	NextSeq 2000 P3 試薬
シングルリード	～4億	～12億
ペアエンド	～8億	～24億

P2試薬	100サイクル	200サイクル	300サイクル
リード長	2 x 50 bp	2 x 75 bp	2 x 150 bp
	1 x 100 bp	2 x 100 bp	

illumina®

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。

RNA-Seqを使用した遺伝子発現プロファイリング（参考リード数）

生物種	トランスクリプトームサイズ (推計)	推奨される最低限のシングルリード数*
単細胞生物 (例、バクテリア、酵母)	0.8 – 3.0 Mb	12万 – 45万リード
多細胞無脊椎動物 (例、線虫、ハエ)	10 – 30 Mb	150万 – 450万リード
複雑な多細胞生物 (例、植物、脊椎動物)	30 – 200 Mb	500万 – 1500万リード

* これらは一般的なガイドラインであり、最初に実験を実施する際の参考です。プロジェクトの統計的要求に満たす十分なカバレッジを得るために最適化を強くお勧めします。

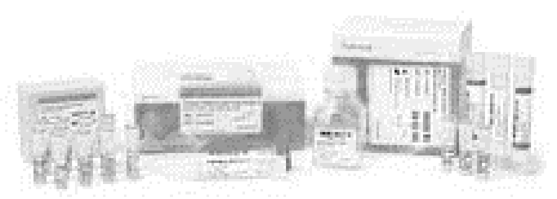
ENCODEプロジェクトには、参考となるRNA-Seqのデータスタンダードがあります。イルミナでは適切な実験計画のため、関連分野と生物種について主要文献を参照することをお勧めします。

RNAシーケンスのワークフロー

RNA抽出

目的の生物からトータルRNAを抽出します。

ライブラリー調製



シーケンス（NGS）

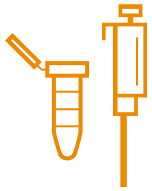
解析

BaseSpace® SEQUENCE HUB

サンプルから結果に至るまでナビゲート

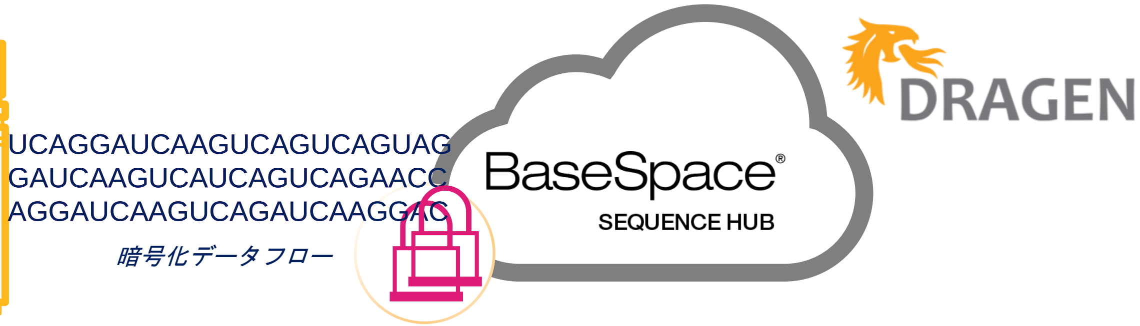
- イルミナのBaseSpace® Suiteは総合的なソリューションを提供

1
ライブラリー調製



2
シーケンス

3
解析



BaseSpace Clarity LIMS

サンプル追跡とワークフロー管理でラボの効率性が向上します。
複数のイルミナシーケンサーがシームレスに一体化されるため、ラン設計と性能QCが容易になります

二次解析

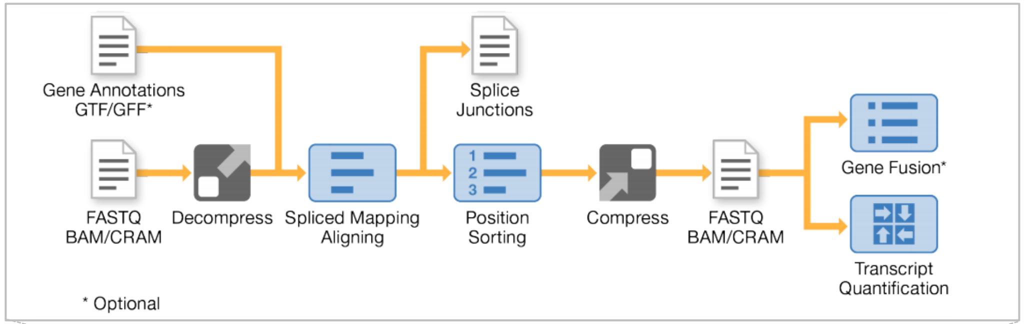
BaseSpace Sequence Hubで、DRAGEN RNAアプリとDRAGEN RNA Differential Expressionアプリを使用してサンプルあたりのコストを抑えて、高精度かつ高速に解析します

三次解析

mRNAの実験では、BaseSpace Correlation Engineで、データを大規模なキュレートナレッジベースと統合することができます

BaseSpace Sequence Hubでデータを安全に格納して直接共有

エンドツーエンドのmRNAシーケンシングワークフロー



Cloud | **Server**

BCL-Convert → DRAGEN RNA Pipeline → DRAGEN RNA Differential Expression Pipeline → Correlation Engine, IGV visualization

一次解析
(.fastq生成)

二次解析

三次解析

DRAGEN™ Bio-IT Platform

超高速でコスト効率の高い、NGS データの二次解析

DRAGENの特徴

Dynamic Read Analysis for GENomics

01 精確

SNPとINDELを高い感度と特異度で検出

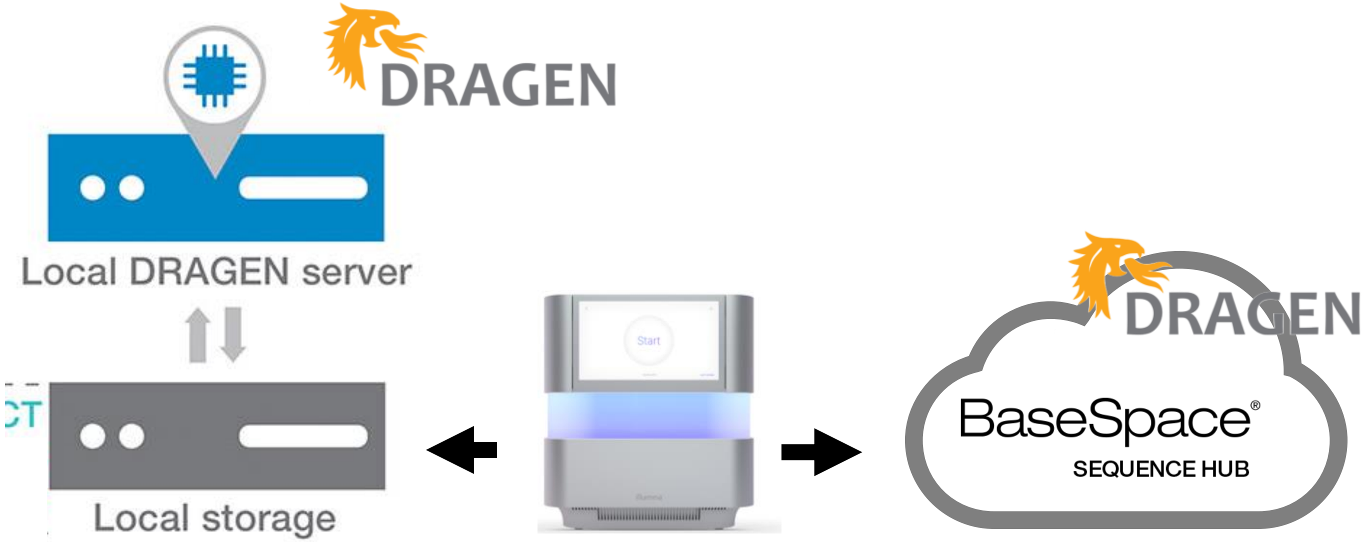
02 超高速

30Xヒト全ゲノム解析所要時間を約25分に短縮

03 コスト効率的でスケーラブル

コストを抑えながら、必要に応じた増強が可能

オンサイトサーバーとクラウドを必要に応じて選べる柔軟性



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。

DRAGEN RNA pipeline とRNA-Seq Alignment 解析時間の比較

サンプル条件: ヒトmRNAサンプル４検体
リード数: 9500万リード、11,000万リード、12,000万リード、13,000万リード
リード長: 2 x 75 bp

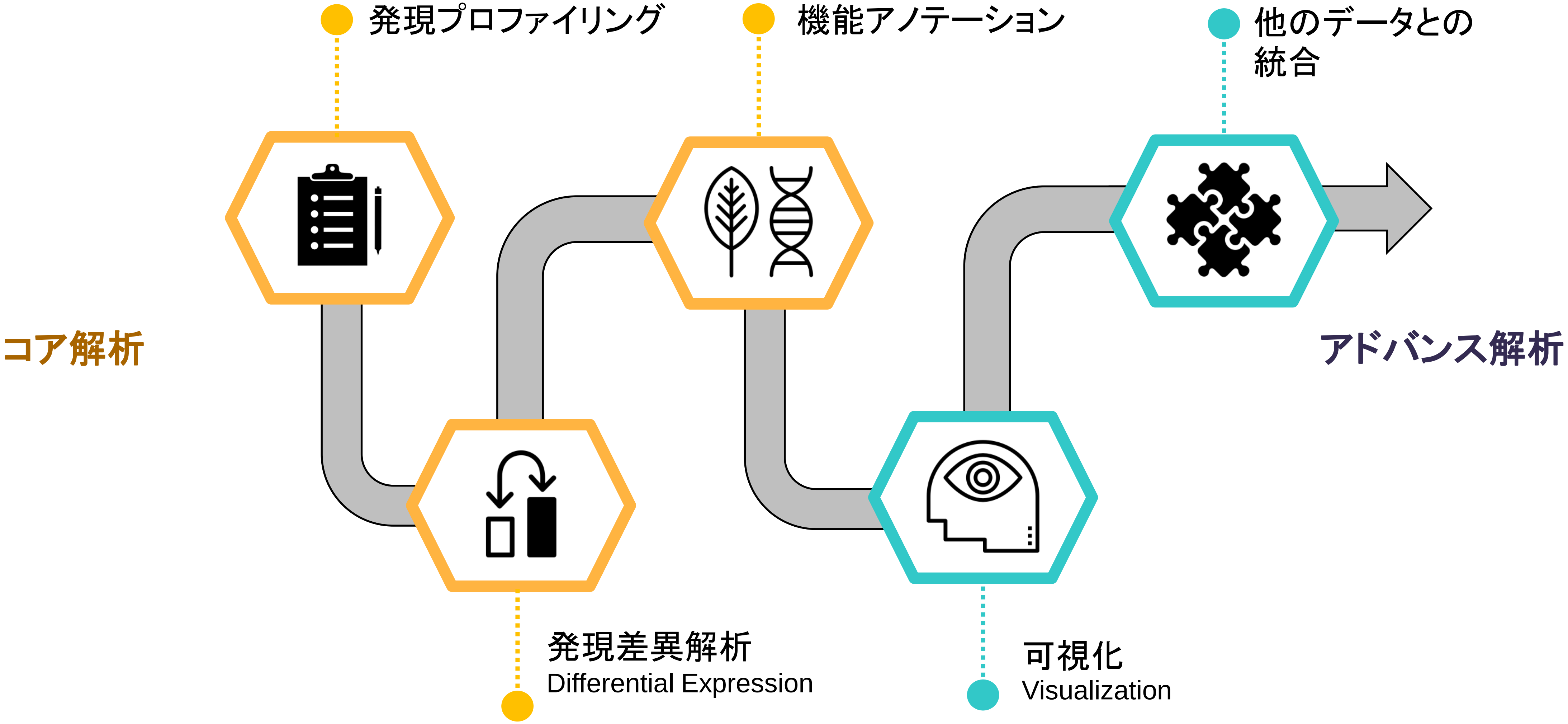
	解析時間	iCredit
RNA-Seq Alignment	4時間 7分 13秒	42.00 iCredit (約4,956 円)
DRAGEN RNA pipeline	**16分 45秒**	5.00 iCredit （約590 円）

DRAGENを使うことにより解析時間、解析コストを
圧倒的に削減することが出来る

illumina®

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。

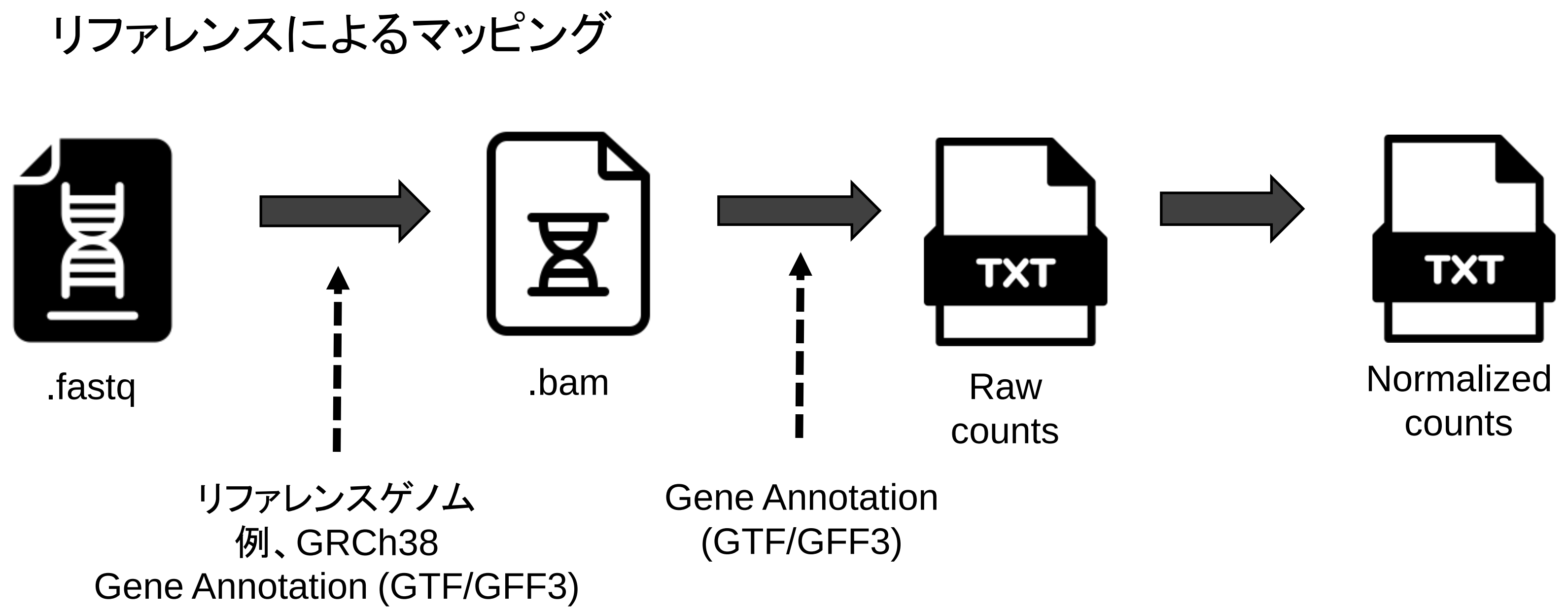
一般的な RNA-seq 解析ワークフロー



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。

RNAサンプルの発現プロファイリング：リードからカウントへ

サンプル中の既知の遺伝子または転写産物の発現プロファイリング



DRAGEN RNA

DRAGEN RNA Pipeline

Input Files

- FASTQ.
- Optional: Custom Reference.
- Optional: Gene Annotation File (GTF, GFF, or GFF3).

Output Files

- Aligned reads - BAM/SAM/CRAM format.
- SJ.out.tab - summarizes the high confidence splice junctions.
- Chimeric.out.junction - information about split-reads which can be used to perform downstream gene fusion detection.
- fusion_candidates.preliminary and .final (if Gene Fusion is enabled) - pre-filtered and post-filtered list of candidate events detected.

DRAGEN

DRAGEN is a Bio-IT Platform that provides ultra-rapid secondary analysis of sequencing data using field-programmable gate array technology (FPGA).

Mapping metrics

Mapping metrics, similar to the metrics computed by the samtools-stats command. Shown on per read group level. To see per-sample level metrics, refer to the general stats table.

Copy table | Configure Columns | Plot | Showing 2/2 rows and 13/78 columns.

Sample Name	M Input reads	Raw depth	Paired	QC-fail	Unmap	Dup	Prop pair	Discord	Singleton
Brain-100ng-Rep1_20549232461	64.5	2.0 x	100.0%	0.00%	0.3%	62.0%	96.8%	2.88%	0.06%
Brain-100ng-Rep1_20549243225	65.0	2.1 x	100.0%	0.00%	0.3%	66.7%	96.8%	2.89%	0.07%

DRAGEN-FastQc

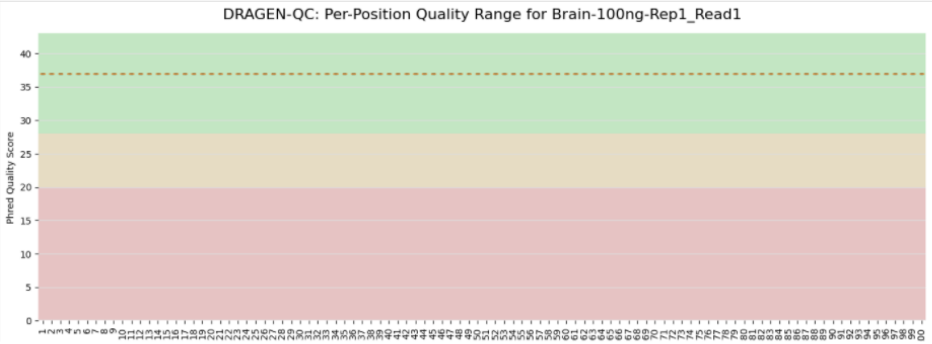
DRAGEN-FastQc is a Bio-IT Platform that provides ultra-rapid secondary analysis of sequencing data using field-programmable gate array technology (FPGA).

Per-Position Quality Score Ranges

The range of quality value across each base position in each sample or read

Flat image plot. Toolbox functions such as highlighting / hiding samples will not work (see the docs).

1 2

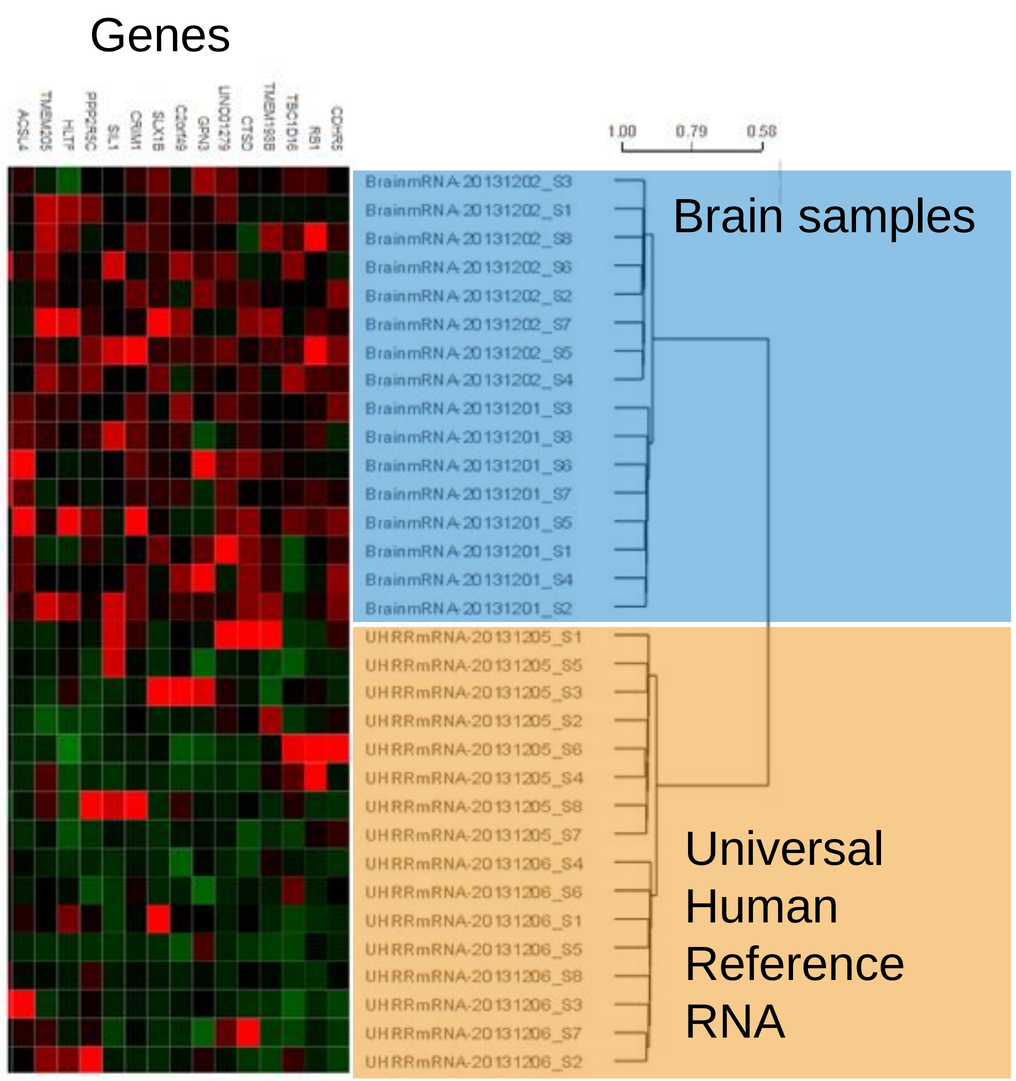


発現差異解析

候補の選択

- 下流解析を行う価値のある候補を絞り込むための指標例
  - 2つの条件下で$\log_2$Fold Changeの大きい遺伝子
  - False discovery rateで調整されたp値

- 発現シグネチャーまたはサブポピュレーションは、サンプル全体の遺伝子発現値のクラスタリングから推測できます。



DRAGEN Differential Expression

DRAGEN Differential Expression

Input Files

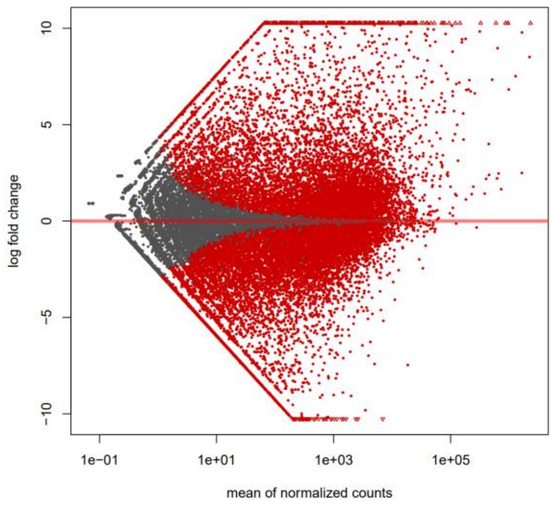
This app consumes Salmon quantification files from the DRAGEN RNA app. Make sure "Enable RNA Quantification" checkbox is selected when running the DRAGEN RNA App. It is also compatible with the RNA-Seq Alignment v2 app, which outputs Salmon quantification files.

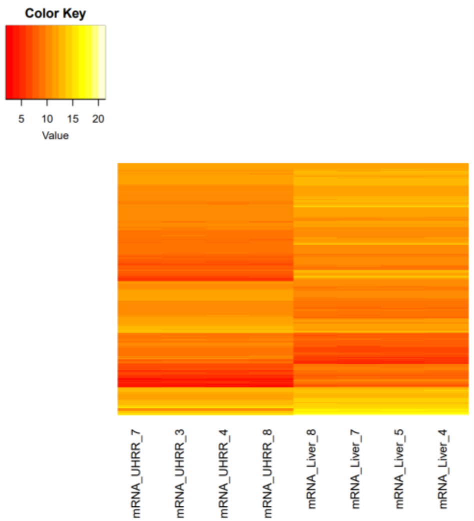
- Salmon quant.sf files
- Select or upload the Gene Annotation File (GTF, GFF, or GFF3) that was used to run the DRAGEN RNA app.

Output Files

The app produces differential expression results produced by DESeq2 at the gene level and transcript level.

- genes.res.csv - CSV file of differentially expressed genes with FDR adjusted p-values





Differential Expression Metrics

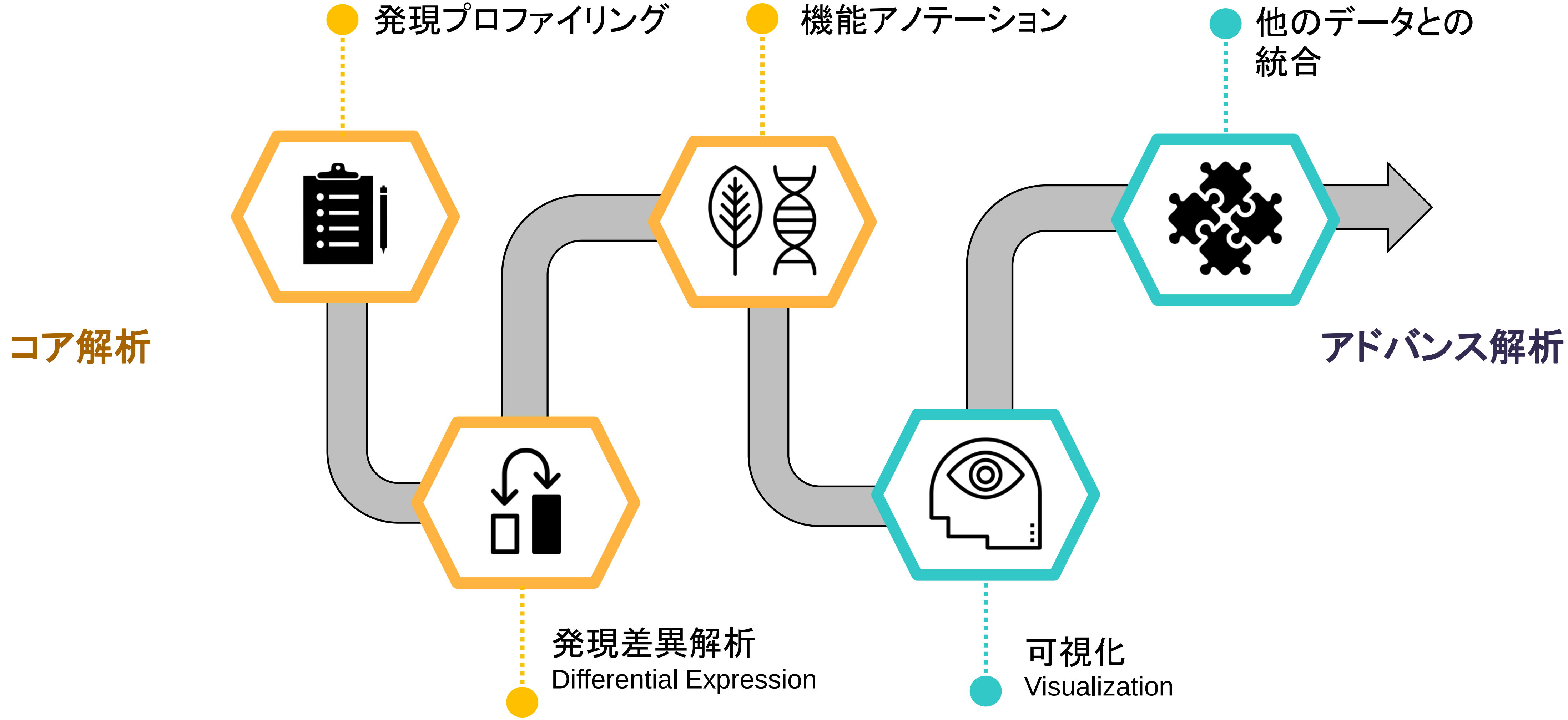
Annotation Gene Count	60609
Assessed Gene Count	20422
Differentially Expressed Gene Count	17715
Merged Gene Counts	Download
Differentially Expressed Genes	Download
Differentially Expressed Transcripts	Download

	mRNA_UHRR_8	mRNA_UHRR_7	mRNA_UHRR_4	mRNA_UHRR_3	mRNA_Liver_4	mRNA_Liver_8	mRNA_Liver_5	mRNA_Liver_7
ENSG00000000003.15	1834	1602	1875	1637	2237	1792	2061	2015
ENSG00000000005.6	22	6	11	18	0	0	0	0
ENSG00000000419.12	2256	1777	2281	1881	316	284	312	283
ENSG00000000457.14	655	581	608	611	347	312	403	347
ENSG00000000460.17	1110	917	1135	939	133	85	98	91
ENSG00000000938.13	502	374	467	442	637	575	606	580
ENSG00000000971.15	209	198	206	195	47644	39244	44943	42544
ENSG00000001036.14	4383	3742	4538	3745	3485	2984	3321	3112

gene_id	baseMean	log2FoldChan	lfcSE	stat	pvalue	padj	stat	control	comparison	gene_name
ENSG00000188641.13	1716.479413	2.609543237	0.069563099	37.51332628	5.59E-308	6.74E-307	OK	483.3040174	2949.654809	DPYD
ENSG00000122786.20	7006.633522	1.647833042	0.043956305	37.48797905	1.45E-307	1.74E-306	OK	3390.066436	10623.20061	CALD1
ENSG00000066583.12	1415.291694	1.647292058	0.043949141	37.48178095	1.82E-307	2.20E-306	OK	684.9647379	2145.61865	ISOC1
ENSG00000137364.5	1269.321715	1.818646024	0.048531552	37.47347766	2.49E-307	3.00E-306	OK	560.7164024	1977.927028	TPMT
ENSG00000178695.6	1546.014585	2.033220512	0.054282932	37.45598175	4.80E-307	5.78E-306	OK	607.0926071	2484.936563	KCTD12
ENSG00000164938.14	1944.896859	2.167177677	0.057873751	37.44664268	6.81E-307	8.20E-306	OK	708.3378068	3181.455911	TP53INP1

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。

一般的な RNA-seq 解析ワークフロー



illumina®

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。

視覚化

目的

- ターゲット遺伝子に関連するイベントを視覚化します。発現変化、融合遺伝子、スプライシング。

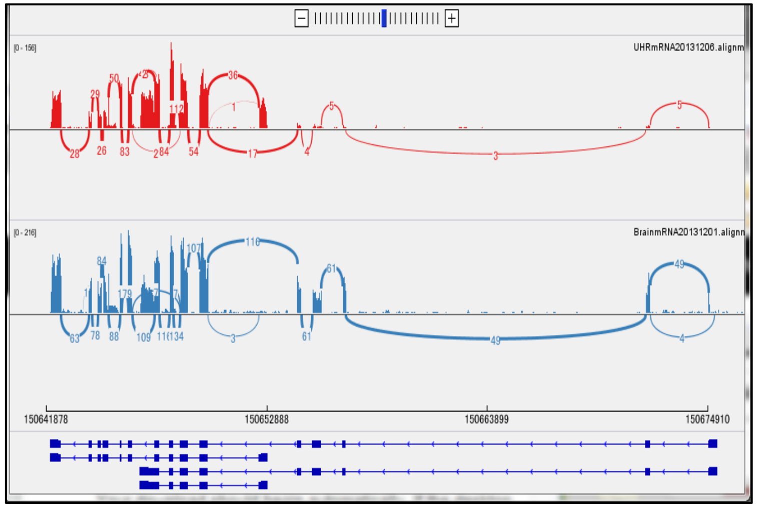
詳細

- インプット: .bam ファイルまたは .bigwig ファイル
- ツール: IGV, Sashimiプロット (スプライシング)
- より高度なユースケース：UCSCゲノムブラウザーを使用したパブリックアノテーションとのオーバーレイ
- 例、ENCODEデータベースに公開されている細胞株データセット、GTExデータベースの正常組織データセット、TCGAデータベースの癌組織データセットの発現やクロマチンプロファイルとオーバーレイ表示します。

視覚化例

発現変化

- IGV

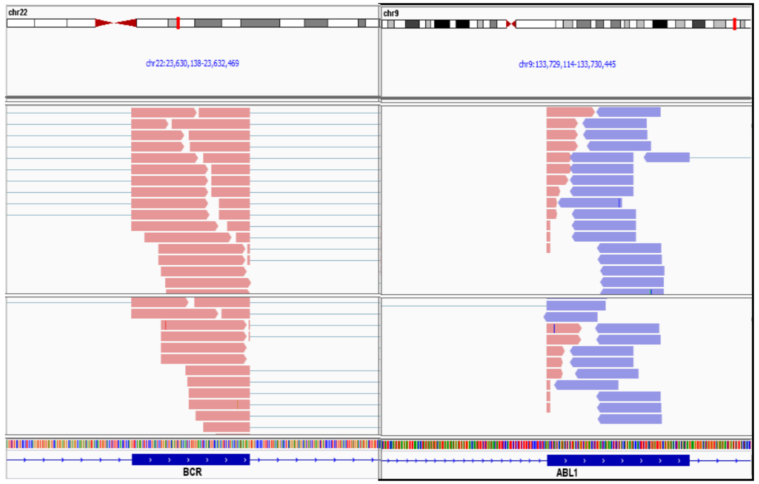


スプライス
アイソフォーム
の検出

- Sashimi Plot

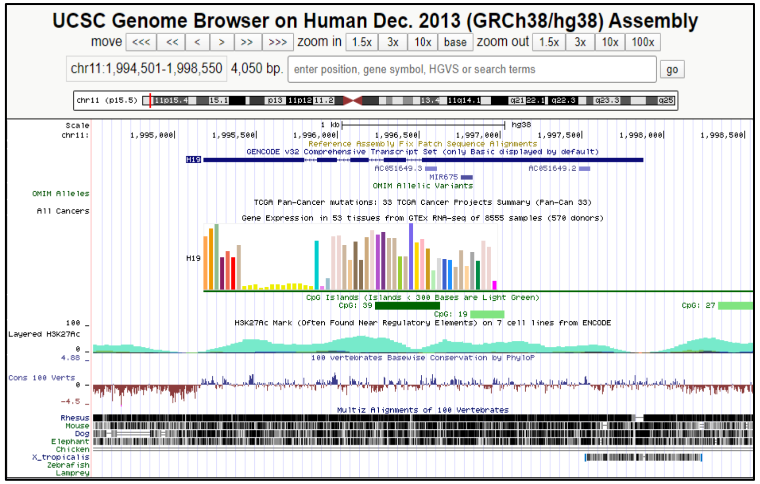
本製品の使用目的は研究に限定されます。診断での使用はできません。

視覚化例



融合遺伝子の検出

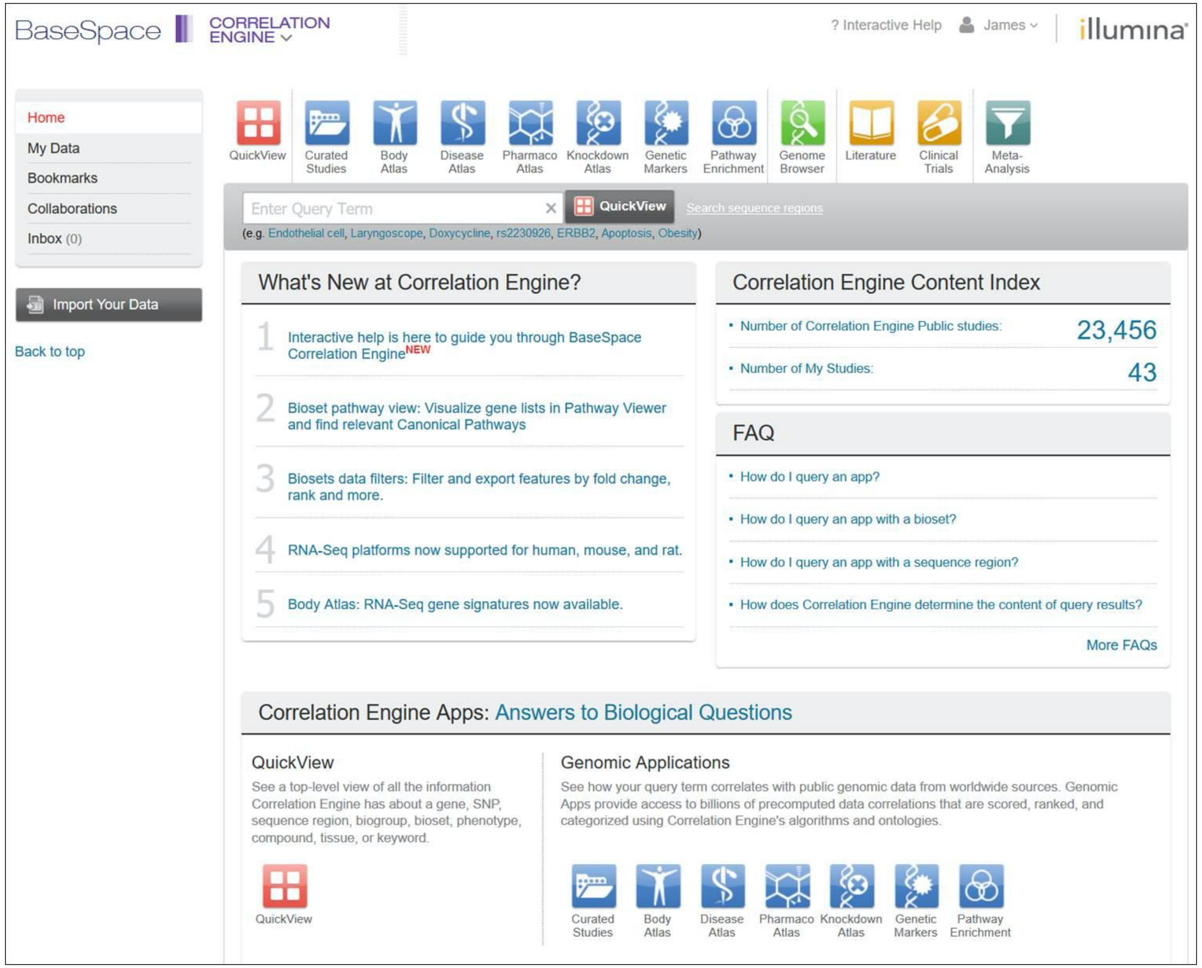
- IGV



パブリックアノテーションとのオーバーレイ

- UCSC Genome Browser

BaseSpace コリレーションエンジン （Correlation Engine）



高品質でキュレートされた公開データの大規模なナレッジベース

- 科学者のインタラクティブなクエリに基づいて、データパターンをリアルタイムで導き出す強力なツール
- データをアップロードすると、過去の23,000件の研究に対してメタ分析を行うシステム
- 組織間でさまざまなデータの共有が可能

まとめ：イルミナのRNAシーケンシングソリューション

	ライブラリー調製		シーケンス		解析
	RNAインプット	Illumina RNA Prep	リード/サンプル[2]	サンプル/ラン[2]	DRAGEN & BaseSpaceオプション[3]
Total RNA コーディング領域と非コーディング領域	1-1000ng[1]	Stranded Total RNA Prep with Ribo-Zero Plus	50M	NextSeq™ 550: 2-8 NextSeq 2000: 8-50 NovaSeq 6000: 16-200	DRAGEN RNA Pipeline DRAGEN Differential Expression
mRNA Seq ポリAテールを持つ転写産物	25-1000ng[1]	Stranded mRNA Prep	25M	NextSeq™ 550: 5-16 NextSeq™ 2000: 16-40 NovaSeq™ 6000: 32-400	DRAGEN RNA Pipeline DRAGEN Differential Expression Correlation Engine (三次解析)
RNA Enrichment w/Exome ターゲット領域	最低10ng	RNA Prep with Enrichment + Exome Panel	25M	NextSeq™ 550: 5-16 NextSeq™ 2000: 16-40 NovaSeq™ 6000: 32-200	DRAGEN RNA Pipeline DRAGEN Differential Expression
（RNA Enrichment w/Respiratory Virus Panel） SARS-CoV-2ゲノム+重複感染の検出	最低10ng	RNA Prep with Enrichment + Respiratory Virus Panel	500k	iSeq™ 100: 8 MiSeq™: 1-50 MiniSeq™: 8-50 NextSeq™ 550: 260-384	DRAGEN RNA Pathogen Detection IDbyDNA Explify RVOP

Clarity LIMS FOR TRACKING & WORKFLOW MANAGEMENT

1. For optimal data quality & FFPE/degraded samples use 10ng minimum
2. Reads per sample can vary based on study goals, number of samples per sequencing run calculated based on listed reads per sample and max 384 UDIs
3. Options for data analysis represented, processing may require data modifications

*

Thank You

ksajiki@illumina.com

illumina®

© 2021 Illumina, Inc. All rights reserved.