

Illumina Protein Prep solutionで、 正確で信頼性の高いタンパク質検出を 達成

DRAGEN™ Protein Quantification二次解析パイプラインを使用したQCステップ
およびデータノーマライゼーションステップのためのガイド

再現性の高い結果

サンプル間、プレート間でも
一貫したタンパク質検出を実現

自動起動解析

DRAGEN二次解析による
高品質な出力の生成

洞察までを統合

Illumina Connected
Multiomicsでノーマライズした
タンパク質量ファイルを解析

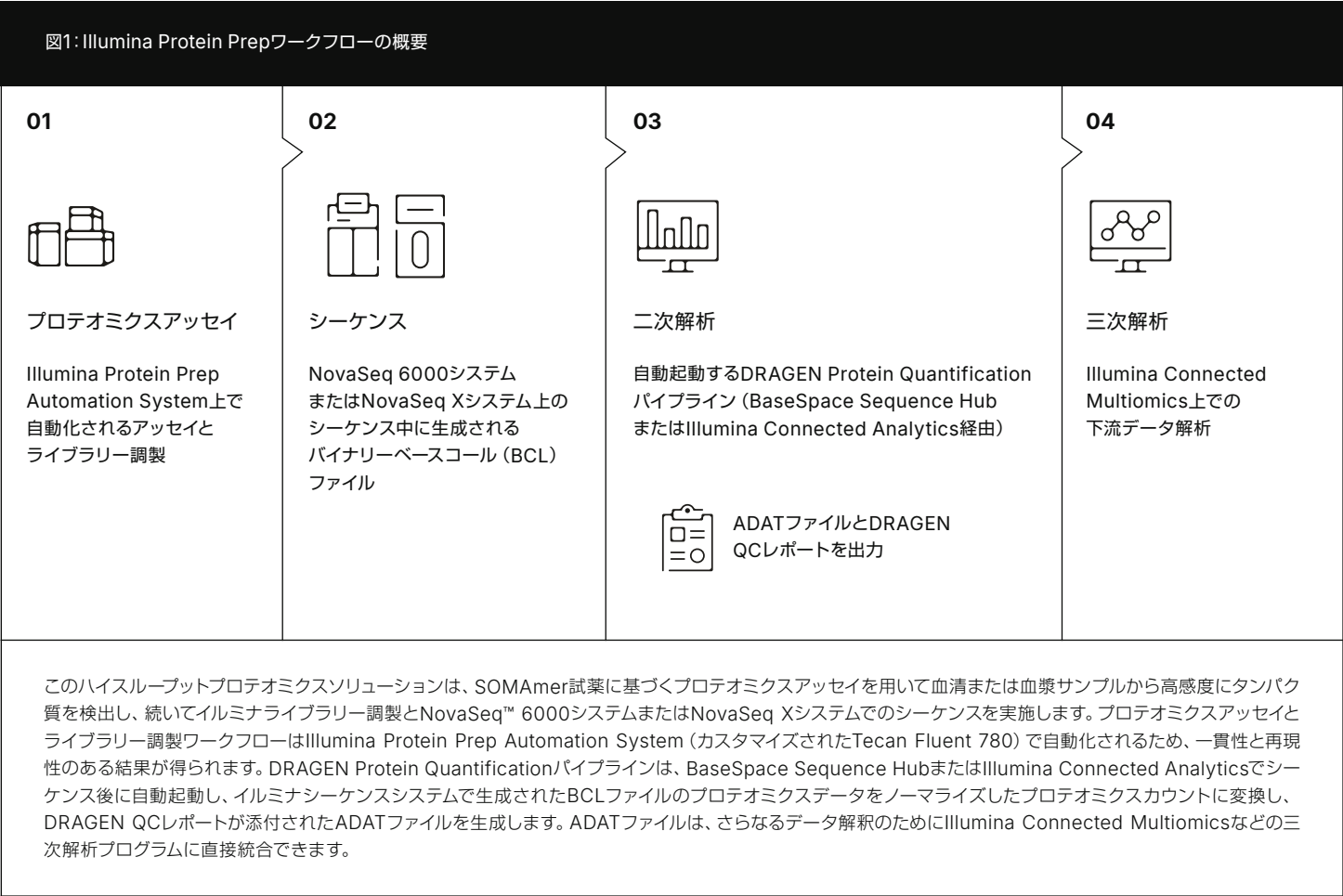
はじめに

タンパク質はヒトの生物学において重要な機能的役割を果たし、健康状態や疾患状態をリアルタイムで反映します。正確で再現性のあるタンパク質定量は、意味のある生物学的洞察のために極めて重要です。革新的なプロテオミクスアッセイであるIllumina Protein Prepは、タンパク質捕捉にSOMAmer® (Slow Off-rate Modified Aptamer) 試薬を使用して、非常に高感度なタンパク質検出を実現します。このアッセイに次世代シーケンサー (NGS) リードアウトとDRAGEN二次解析のパイオインフォマティクスの力を組み合わせることで、効率的かつハイスループットなサンプルから結果までのプロテオミクスソリューションを提供します。

Illumina Protein Prepは単一サンプル中の数千のタンパク質を同時に検出する、バイオマーカー探索のためのパワフルなプラットフォームです。しかし、プレート間のバイアスやハイスループットプロテオミクスの技術的なばらつきにより、特に効果量が小さいバイオマーカーでは、発現差異が不明瞭になることがあります。DRAGEN Protein Quantificationパイプラインは、ロバストなノーマライゼーション法を用いて二次解析中にこの課題に対処し、実験間で高

品質な比較データを生成します。NovaSeq™ 6000システムまたはNovaSeq Xシステム上でシーケンスが完了した後、パイプラインは自動的にクラウドで起動します。DRAGEN QCレポートに加え、Illumina Connected Multiomicsを用いた下流の三次解析に使用できるノーマライズしたタンパク質量を記載したADATファイルを生成し (図1)、プロテオームに関する詳細な生物学的洞察を提供します。ローカル解析オプションが必要な場合は、DRAGEN Protein QuantificationパイプラインをDRAGENオンプレミスサーバーにインストールすることができます。

このテクニカルノートでは、サンプルおよび技術的なばらつきを除去した高品質なプロテオミクスデータを作成する、DRAGEN Protein Quantificationパイプラインに組み込まれた品質管理 (QC) とノーマライゼーションステップについて説明します。



データノーマライゼーションが必要である理由

技術的ばらつきは、あらゆるハイスループットのプロテオミクスアッセイに固有の性質であり、サンプル調製、内因性のサンプル不均一性、ハイブリダイゼーション効率、PCR効率、シーケンスなど、複数の原因から生じる可能性があります。サンプル内の技術的ばらつきに加え、オペレーター、装置、環境効果の違いにより、同一プレート上のすべてのサンプルに影響を及ぼすバッチ効果も生じる可能性があります。効果的なノーマライゼーションまたはバッチ補正がなければ、このような系統的なばらつきのために真の生物学的相違が不明瞭になる可能性があり、不正確な結論につながる可能性があります。

データを標準化するデータノーマライゼーションは、このばらつきの修正において非常に重要であり、サンプル、プレート、さらに異なるアッセイのランにわたって意味のある比較を可能にします。DRAGEN Protein Quantificationパイプラインを使った包括的なノーマライゼーション戦略を実施することで、信頼性と再現性に優れたタンパク質定量データを得ることができます。

Illumina Protein Prepのコントロール

Illumina Protein Prepアッセイには、アッセイのばらつきを軽減するため、プレートコントロールと各サンプル中のSOMAmer試薬コントロールが含まれています。どちらの種類のコントロールも、DRAGEN Protein Quantificationパイプラインによるノーマライズしたタンパク質発現値の生成に使用します。

プレートコントロール

各96ウェルプレートには、85サンプルと、ブランク、キャリブレーターコントロールおよびQCサンプルを含む11のプレートコントロールを入れます（図2）。これらのコントロールはIllumina Protein Prepアッセイに含まれており、ノーマライゼーションおよびQC用に血漿または血清サンプルとともに各プレートに入れてランを実施します。これらは、技術的ばらつきを検出して修正するための基準となり、異なるプレート上でランしたサンプルを揃える際のベースになります。

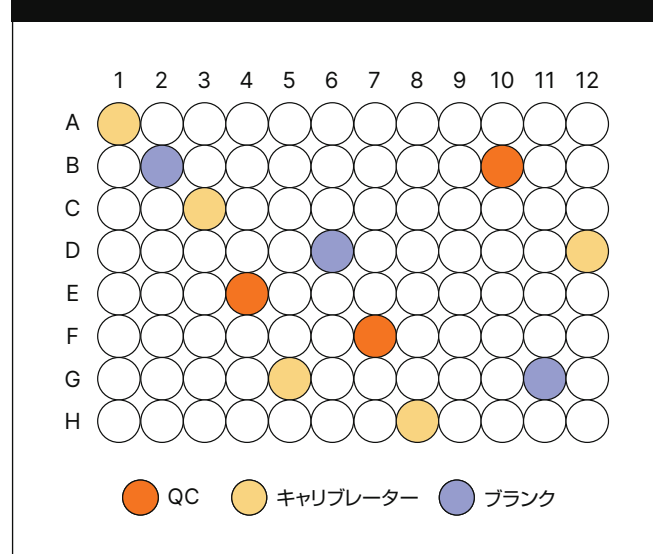
ブランク

ブランクサンプルは何も添加されていないバッファーであり、Illumina Protein Prepアッセイの非特異的バックグラウンドの推定に役立ちます。これらのコントロールは、血漿または血清サンプルよりもカウント数が非常に低くなるはずで、ブランクウェルのカウント数が高い場合、ランの非特異的バックグラウンドが強い、あるいはサンプルにコンタミネーションが生じたことを示します。

キャリブレーターコントロール

キャリブレーターコントロールはプレート間のばらつきの修正に役立ちます。キャリブレーターコントロールには、解析するサンプルタイプに応じて、ヒトのプール血漿または血清のいずれかを使用します。プレートキャリブレーションノーマライゼーションは、SOMAmer試薬ごとのキャリブレーターコントロールの中央値と、対応するキャリブレーションリファレンスを比較することで実施します。

図2：コントロールを含むプレートレイアウトの例。



QCサンプル

QCサンプルは品質管理のチェックとして働き、プレートが期待どおりに反応したかどうかを判定します。これらのコントロールサンプルには、解析するサンプルタイプに応じて、ヒトのプール血漿または血清のいずれかを使用します。プレート品質は、SOMAmer試薬ごとのこれらのサンプルの中央値と対応するQCリファレンスを比較することで評価します。

SOMAmer試薬コントロール

SOMAmer試薬コントロールは、ハイブリダイゼーション前に既知濃度で各サンプルに組み込まれており、サンプルごとのスケール係数を計算し、サンプルをノーマライゼーションし、品質を評価します。

データノーマライゼーションステップ

プロテオミクスカウントの生データのノーマライゼーションには、Illumina Protein Prepワークフロー中に生じるサンプルとプレートの系統的なバイアスを除去するための、複数のステップが必要です。これらの調整は、ハイブリダイゼーションノーマライゼーション、外部リファレンス中央値ノーマライゼーション、キャリブレーション、プレートスケーリングの4つの不連続なステップにわたって実施されます (図3)。

ハイブリダイゼーションノーマライゼーション

このステップは、アッセイのハイブリダイゼーションおよびライブラリー調製ステップ中に生じるサンプル間ばらつきを修正します。サンプル固有のスケール係数は、各サンプルのハイブリダイゼーションノーマライゼーションコントロールのカウント数と、プレート内のすべてのサンプルの中央値を比較して計算します。

外部リファレンス中央値ノーマライゼーション

このステップはコントロール用以外のサンプルの総タンパク質存在量測定値の差を修正します。スケール係数は、観測されたタンパク質測定値と、各タンパク質の予測値のリファレンスを比較して計算します。

キャリブレーション

キャリブレーションは、プレートコントロールセクションで概説した5つのキャリブレーターサンプルを使って実施し、個々のSOMAmer試薬に影響を及ぼすバッチ効果を補正します。このステップにより、異なるプレートやランから生成されたデータ間の正確な比較ができます。各SOMAmer試薬にキャリブレーションを実施し、キャリブレーターとキャリブレーターコントロールの予測リファレンス値のばらつきを補正します。

図3：DRAGEN Protein Quantificationパイプラインで実施したノーマライゼーションステップの概要。

01	02	03	04	05	06	07
サンプル調製	結合	SOMAscan アッセイ	ハイブリダイ ゼーション	キャプチャーと 洗浄	ポストPCRの クリーンアップ	シーケンス
1.5時間	3.5時間	2.3時間	19時間	3.3時間	0.5時間	シーケンスラン時間 はサンプル数、シス テム、フローセルに よって決まります
外部リファレンス中央値のノーマライゼーション			ハイブリダイゼーションのノーマライゼーション			
プレートスケーリングとキャリブレーション						

キャリブレーションは以下の2つのステップで実施します。

- ステップ1: キャリブレーター中央値とシーケンサー由来リファレンスを比較し、それに応じて各SOMAmer試薬を調整します。
- ステップ2: 更新されたキャリブレーター中央値とシーケンサー由来リファレンスを比較することで、異なる装置で生成されたデータを比較することができます。キャリブレーターの中央値とリファレンス値のアライメントに使用したスケール係数は、そのプレート上のすべてのサンプルに適用されます。

プレートスケールリング

プレートスケールリングは、キャリブレーター全体の各SOMAmer試薬の測定値の中央値を用いて、これらを外部キャリブレーションリファレンスとアラインメントすることで、プレート間の総タンパク質測定値のばらつきに対処します。

スケールリングは以下の2つのステップで実施します。

1. ステップ1: NovaSeq 6000システムやNovaSeq Xシステムなどのシーケンス装置由来のリファレンスに基づいて測定値を調整します。
2. ステップ2: シーケンサーを超えて残るリファレンスを用いてこの調整の精度を高め、異なる装置で生成されたデータ間の比較を可能にします。DRAGEN Protein Quantificationレポートには、スケールリングのステップ1から得られたQCメトリクスのみが表示されます。

両端のQCチェック率

QCステップは、各プレートのQCコントロールを使用してノーマライゼーションの最後に実施します。このステップでは、3つのQCサンプルのレプリケートにわたる各SOMAmer試薬測定値の中央値を外部QCリファレンスと比較し、SOMAmer試薬に固有のスケール係数とQCメトリクス (QCCheckTailPercent) を割り当てます。

DRAGEN Protein Quantificationパイプラインで実施されたさまざまなノーマライゼーション法は、Illumina Protein Prepワークフローの異なるステージから生じた関連性のないばらつきを最小限に抑えます (図4)。

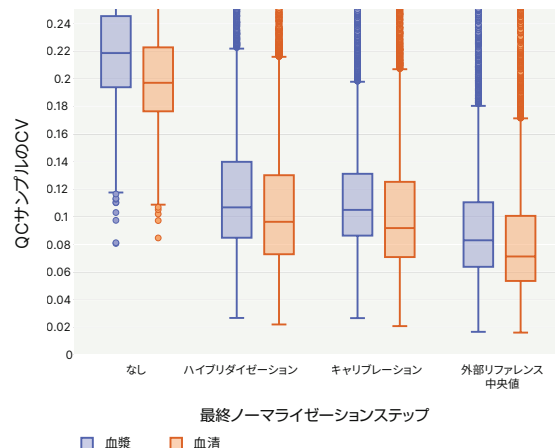
DRAGEN Protein Quantificationの出力

DRAGEN Protein QuantificationパイプラインはIllumina Protein Prepデータの解析に重要な2種類のファイルを出力します。1つ目は各ノーマライゼーションステップで得られた一連のADATファイル、2つ目はランのQCを要約したDRAGEN QCレポートです。

ADATファイル

DRAGEN Protein Quantificationパイプラインはカウントの生データを処理し、実施した各ノーマライゼーションステップの出力として、adat拡張子の付いたタブ区切りASCIIテキストファイルを生成します。ADATファイルは一連のサンプル (行) にわたる一連の分析物 (列) の測定値および分析物とサンプルの説明情報を含みます。

図4: データのばらつきに対するノーマライゼーションの影響。



アッセイの生データのノーマライゼーションはプレート内の変動係数 (CV) の低下につながります。データは、血漿および血清のQCサンプルを用いたIllumina Protein Prep 6Kアッセイを用いて生成しました。中央値CV5.5%未満のIllumina Protein Prep 9.5Kアッセイがご利用いただけるようになりました。

a. 健康なドナーサンプルを用いて計算された中央値CVの期待値。

図5：ADATファイル構造。



ADATファイルは以下の構造からなるタブ区切りファイルです (図5)。

- **ランメタデータ**はラン全体または個々のプレートに関連するメトリクスを含みます (例：RunIDおよび両端に含まれるQCチェック率と関連するメトリクス)
- **SOMAmerメタデータ**は各SOMAmer試薬またはタンパク質ターゲットに関連するメトリクスを含みます (例：SOMAmer ID、および特定の遺伝子とターゲットID)
- **サンプルメタデータ**は各サンプルに関連するメトリクスを含みます (例：Sample ID、Plate ID、QC Metrics)
- **SOMAmer数値**は各サンプル中の各SOMAmerの相対存在量を反映するNGSカウントのマトリクスです

このフォーマットはサンプル数、分析物の数と種類、サンプルの説明を柔軟に提供できるようにデザインされており、生物学的洞察を得るための下流の三次解析に使用できます。

DRAGEN QCレポート

DRAGEN Protein Quantification QCレポートはADATファイルと一緒に生成される、各個別サンプルとプレート全体のQCのサマリーです。個別サンプルのQCはハイブリダイゼーションノーマライゼーションステップおよび外部リファレンス中央値ノーマライゼーションステップに由来します。プレートQCはキャリブレーションノーマライゼーションステップおよびQCステップに由来します。

三次解析用のADATファイルの使用

ノーマライズしたADATファイルはIllumina Connected Analytics内のIllumina Connected Multiomicsを用いたさらなるデータ解析にそのまま使うことができます。Illumina Connected Multiomicsはデータ管理と選択のためのグラフィカルユーザーインターフェースを備えており、タンパク質発現差異、パスウェイ解析、遺伝子セット濃縮解析など、高度なデータ解釈に使用できます。ADATファイルのコンテンツはR (SomadataIO) およびPython (Canopy) で提供されるオープンソースパッケージを使用して解析することもできます。

まとめ

プロテオミクスの生データのノーマライゼーションは、Illumina Protein Prepワークフロー中に生じるアッセイとサンプルのバイアスを最小限に抑えるために非常に重要です。DRAGEN Protein Quantificationパイプラインは、複数のノーマライゼーションステップとキャリブレーションステップを用いて系統的なばらつきを抑え、真の生物学的なばらつきが除去されるリスクを最小化し、より正確な結果を提供します。このパイプラインはクラウドまたはオンプレミスサーバーで利用できます。DRAGEN Protein Quantificationパイプラインは、NovaSeq 6000システムまたはNovaSeq Xシステム上でシーケンスが完了すると、自動的にクラウドで起動し、DRAGEN QCレポートとノーマライズしたタンパク質量を含むADATファイルを生成します。ADATファイル出力はIllumina Connected Multiomicsを用いた下流三次解析にそのまま使うことができ、プロテオームに関するさらに深い生物学的洞察が得られます。

詳細はこちら →

[Illumina Protein Prep](#)

[DRAGEN二次解析](#)

[Illumina Connected Analytics](#)

イルミナ株式会社

〒108-0014 東京都港区芝 5-36-7 三田ベルジュビル 22 階

Tel (03) 4578-2800 Fax (03) 4578-2810

jp.illumina.com

 www.facebook.com/illumina

販売店

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。 販売条件 : jp.illumina.com/tc

© 2025 Illumina, Inc. All rights reserved.

すべての商標および登録商標は、Illumina, Inc. または各所有者に帰属します。

商標および登録商標の詳細は jp.illumina.com/company/legal.html をご覧ください。

予告なしに仕様および希望販売価格を変更する場合があります。

illumina[®]