

Obtenha detecção precisa e confiável de proteínas com a solução Illumina Protein Prep



Um guia para as etapas de CQ e normalização de dados usando o pipeline de análise secundária do DRAGEN™ Protein Quantification

Resultados reproduzíveis

Obtenha detecção consistente de proteínas em amostras e placas

Análise de inicialização automática

Gere resultados de alta qualidade com o DRAGEN secondary analysis

Insights integrados

Analise arquivos de contagem de proteínas normalizados facilmente com o Illumina Connected Multiomics

Introdução

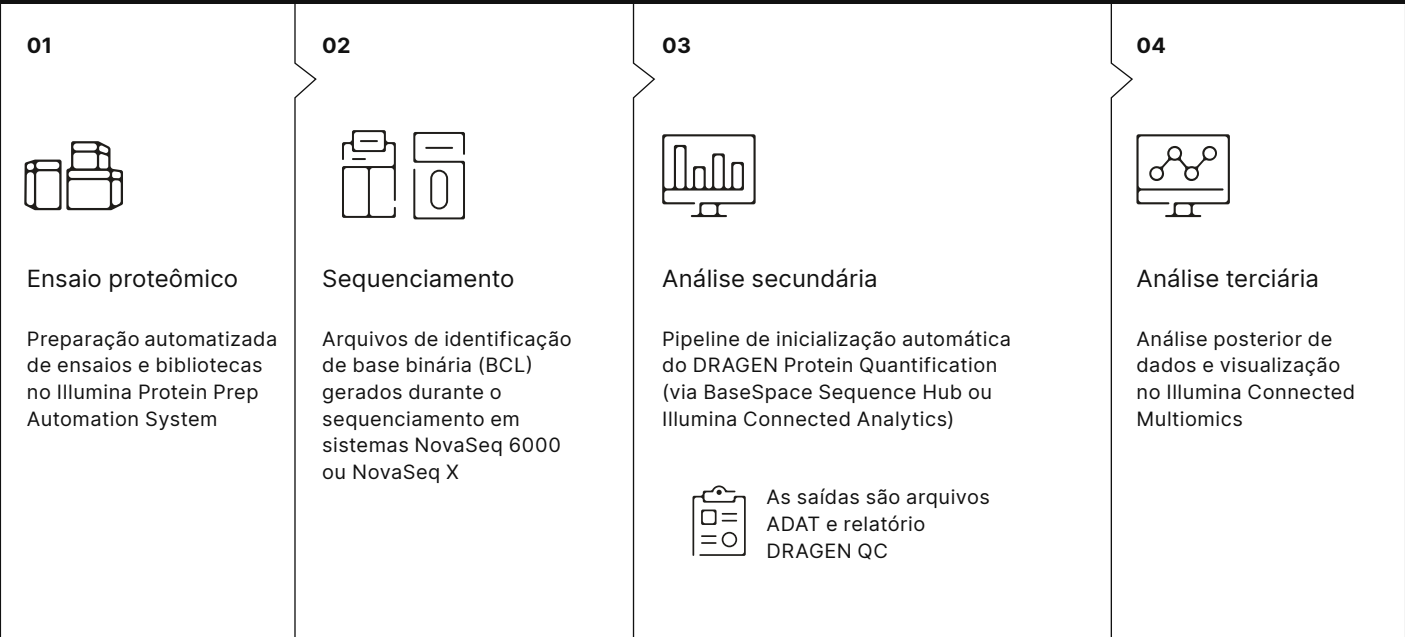
As proteínas desempenham um papel fundamental na biologia humana, refletindo um panorama em tempo real dos estados de saúde e doenças. A quantificação precisa e reproduzível de proteínas é crucial para insights biológicos significativos. O Illumina Protein Prep é um ensaio proteômico inovador que usa SOMAmer® Reagents (aptâmero modificado de baixa velocidade) para captura de proteínas, permitindo a detecção de proteínas altamente sensíveis. A combinação desse ensaio com a leitura de sequenciamento de última geração (NGS) e o poder de bioinformática do DRAGEN secondary analysis fornece uma solução proteômica simplificada e de alto rendimento de amostra para resultados.

O Illumina Protein Prep detecta milhares de proteínas simultaneamente em uma única amostra, fornecendo uma plataforma poderosa para a descoberta de biomarcadores. No entanto, o viés placa a placa e a variabilidade técnica de plataformas proteômicas de alto rendimento podem obscurecer a expressão

diferencial, particularmente para biomarcadores com tamanhos de efeito pequenos. O pipeline do DRAGEN Protein Quantification aborda esse desafio durante a análise secundária, com métodos completos de normalização para fornecer dados comparáveis de alta qualidade em experimentos. Na nuvem, o pipeline é iniciado automaticamente após a conclusão do sequenciamento em um sistema NovaSeq™ 6000 ou NovaSeq X e gera um relatório DRAGEN QC e arquivos ADAT com contagens de proteínas normalizadas que podem ser usadas para análise terciária posterior com o Illumina Connected Multiomics (Figura 1) para obter informações biológicas mais profundas sobre o proteoma. O pipeline do DRAGEN Protein Quantification pode ser instalado em um servidor local do DRAGEN para usuários que precisam de uma opção de análise local.

Esta nota técnica explica as etapas de controle de qualidade (CQ) e normalização incorporadas no pipeline do DRAGEN Protein Quantification para eliminar a variabilidade técnica e de amostras para dados proteômicos de alta qualidade.

Figura 1: Uma visão geral do fluxo de trabalho do Illumina Protein Prep.



A solução proteômica de alto rendimento usa um ensaio proteômico baseado em SOMAmer Reagent para detecção de proteínas sensíveis de amostras de soro ou plasma, seguido pela preparação e sequenciamento da biblioteca da Illumina nos sistemas NovaSeq 6000 ou NovaSeq X. O ensaio de proteômica e o fluxo de trabalho de preparação da biblioteca são automatizados no Illumina Protein Prep Automation System, um Tecan Fluent 780 personalizado, para resultados consistentes e reproduzíveis. O pipeline do DRAGEN Protein Quantification é iniciado automaticamente após o sequenciamento no BaseSpace Sequence Hub ou no Illumina Connected Analytics, transformando os dados de proteômica em arquivos BCL produzidos pelos sistemas de sequenciamento da Illumina em contagens de proteômica normalizadas em um arquivo ADAT com um respectivo relatório de CQ do DRAGEN. O arquivo ADAT pode ser integrado diretamente em programas de análise terciária, incluindo o Illumina Connected Multiomics, para interpretação adicional dos dados.

Por que a normalização de dados é necessária?

A variabilidade técnica é inerente a qualquer ensaio proteômico de alto rendimento e pode surgir de várias fontes, incluindo preparação de amostras, heterogeneidade intrínseca de amostras, eficiência de hibridização, eficiência de PCR e sequenciamento. Além da variabilidade técnica em uma amostra, podem ocorrer efeitos em lote que afetam todas as amostras na mesma placa, devido a diferenças nos efeitos ambientais, do operador ou do instrumento. Sem normalização eficaz ou correção de lotes, essas variações sistemáticas podem obscurecer diferenças biológicas verdadeiras, levando a conclusões imprecisas.

A normalização de dados é essencial para corrigir essa variabilidade, padronizar dados e permitir comparações significativas entre amostras, placas e até mesmo diferentes corridas de ensaio. A implementação de estratégias abrangentes de normalização com o pipeline do DRAGEN Protein Quantification fornece dados confiáveis e reproduzíveis de quantificação de proteínas.

Controles do Illumina Protein Prep

O ensaio Illumina Protein Prep inclui controles de placas e controles de SOMAmer Reagent em cada amostra para mitigar a variabilidade do ensaio. Ambos os tipos de controles são usados para gerar contagens de expressão de proteína normalizadas com o pipeline do DRAGEN Protein Quantification.

Controles da placa

Cada placa de 96 poços acomoda 85 amostras e 11 controles de placa (Figura 2), incluindo amostras em branco, controles de calibrador e amostras de CQ. Esses controles estão incluídos no ensaio Illumina Protein Prep e são executados em cada placa, juntamente com as amostras de plasma ou soro para normalização e CQ. Eles servem como referências para detectar e corrigir desvios técnicos e fornecem uma base para alinhar amostras executadas em placas diferentes.

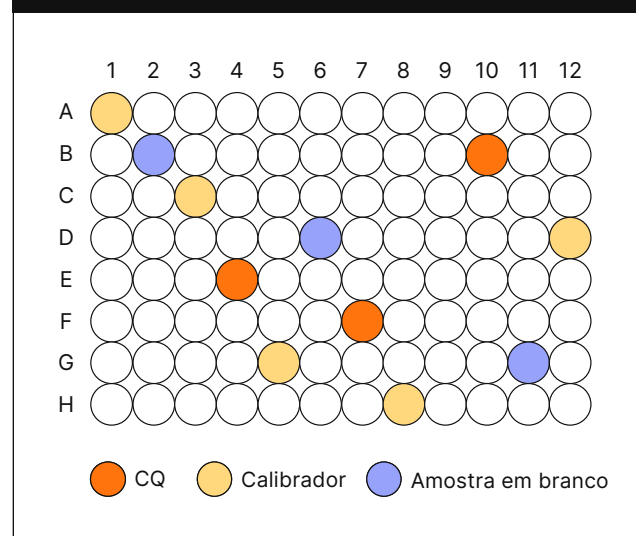
Amostras em branco

As amostras em branco consistem em solução tampão pura e fornecem uma estimativa de fundo não específico no ensaio do Illumina Protein Prep. A expectativa é que esses controles tenham contagens significativamente menores do que as amostras de plasma ou soro. Contagens elevadas em poços em branco podem indicar que uma corrida tem fundo elevado ou que ocorreu contaminação da amostra.

Controles do calibrador

Os controles do calibrador permitem a correção da variabilidade placa a placa. Esses controles são plasma ou soro humano agrupado, dependendo dos tipos de amostra que estão sendo executados. A normalização da calibração da placa é realizada comparando a mediana dos controles do calibrador por SOMAmer Reagent com a referência de calibração correspondente.

Figura 2: Exemplo de layout de placa com controles.



Amostras de CQ

As amostras de CQ fornecem uma verificação de controle de qualidade para determinar se a placa teve o desempenho esperado. Essas amostras de controle são plasma ou soro humano agrupado, dependendo dos tipos de amostra que estão sendo executados. A qualidade da placa é avaliada pela comparação da mediana dessas amostras por SOMAmer Reagent com a referência de CQ correspondente.

Controles do SOMAmer Reagent

Os controles do SOMAmer Reagent são incorporados em cada amostra individual em concentrações conhecidas antes da hibridização e permitem a computação de fatores de escala específicos da amostra para normalizar as amostras e avaliar a qualidade.

Etapas de normalização de dados

A normalização das contagens de proteômica bruta envolve várias etapas para remover vieses sistemáticos em amostras e placas que podem surgir durante o fluxo de trabalho do Illumina Protein Prep. Esses ajustes são realizados em quatro etapas não consecutivas: normalização da hibridização, normalização mediana de referência externa, calibração e dimensionamento de placas (Figura 3).

Normalização da hibridização

Esta etapa corrige a variabilidade de uma amostra para outra que pode ocorrer durante os estágios de hibridização e preparação da biblioteca do ensaio. Os fatores de escala específicos da amostra são calculados pela comparação das contagens dos controles de normalização de hibridização em cada amostra com seus valores medianos em todas as amostras na placa.

Normalização mediana de referência externa

Esta etapa corrige as diferenças na medição da abundância total de proteínas das amostras sem controle. Um fator de escala é calculado comparando as medições de proteína observadas com uma referência dos valores esperados para cada proteína.

Calibração

A calibração é realizada usando as cinco amostras do calibrador descritas na seção de controles da placa e corrige os efeitos do lote que afetam determinados SOMAmer Reagents. Esta etapa permite comparações precisas entre dados produzidos de diferentes placas e corridas. Para cada SOMAmer Reagent, a calibração corrige a variação entre os calibradores e os valores de referência esperados para o controle do calibrador.

Figura 3: Visão geral das etapas de normalização realizadas pelo pipeline do DRAGEN Protein Quantification.

01	02	03	04	05	06	07
Preparação da amostra	Ligação	Ensaio SOMAscan	Hibridização	Captura e limpeza	Limpeza pós-PCR	Sequenciamento
1,5 h	3,5 h	2,3 h	19 h	3,3 h	0,5 h	Os tempos de corrida do sequenciamento dependem do número de amostras, do sistema e da lâmina de fluxo
Normalização mediana de referência externa			Normalização da hibridização			
Dimensionamento e calibração da placa						

A calibração é realizada em duas etapas:

- Etapa 1: Compara as medianas do calibrador com uma referência derivada do sistema de sequenciamento e ajusta cada SOMAmer Reagent de acordo.
- Etapa 2: Compara as medianas atualizadas do calibrador com uma referência derivada de sistemas de sequenciamento, permitindo comparações entre dados gerados em diferentes instrumentos. O fator de escala usado para alinhar a mediana dos calibradores ao valor de referência é aplicado a todas as amostras na placa.

Dimensionamento da placa

O dimensionamento da placa aborda variações nas medições totais de proteína entre as placas usando a mediana de cada medição de SOMAmer Reagent nos calibradores e alinhando-as a uma referência de calibração externa.

O processo de dimensionamento ocorre em duas etapas:

- Etapa 1: Ajusta as medições com base em uma referência derivada do instrumento de sequenciamento, ou seja, sistema NovaSeq 6000 ou NovaSeq X.
- Etapa 2: Refina o ajuste usando uma referência que permanece nos sistemas de sequenciamento, permitindo comparações entre dados gerados em diferentes instrumentos. As métricas de CQ somente da primeira etapa de dimensionamento são exibidas no relatório do DRAGEN Protein Quantification.

Percentual de verificação de CQ nos extremos

Uma etapa de CQ é realizada no final da normalização usando os controles de CQ em cada placa. Esta etapa compara a mediana de cada medição do SOMAmer Reagent entre as três réplicas de amostras de CQ com uma referência de CQ externa para atribuir um fator de escala específico do SOMAmer Reagent e uma métrica de CQ (QCCheckTailPercent).

Os vários procedimentos de normalização implementados no pipeline do DRAGEN Protein Quantification minimizam a variabilidade estranha originada de estágios distintos do fluxo de trabalho do Illumina Protein Prep (Figura 4).

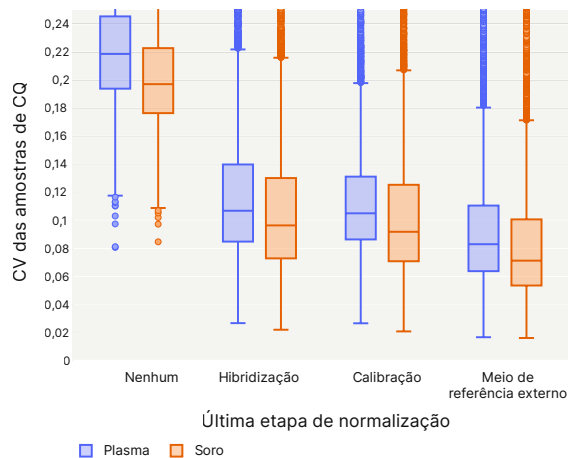
Resultado do DRAGEN Protein Quantification

O pipeline do DRAGEN Protein Quantification produz dois tipos de arquivos que são críticos para analisar os dados do Illumina Protein Prep. O primeiro é uma série de arquivos ADAT de cada etapa de normalização e o segundo é um relatório DRAGEN QC resumindo o CQ da corrida.

Arquivos ADAT

O pipeline do DRAGEN Protein Quantification processa contagens brutas e gera um arquivo de texto ASCII delimitado por tabulação com uma extensão ADAT como saída para cada etapa de normalização realizada. O arquivo ADAT contém medições para uma série de analitos (colunas) em uma série de amostras (linhas) e inclui a descrição do analito e as informações da descrição da amostra.

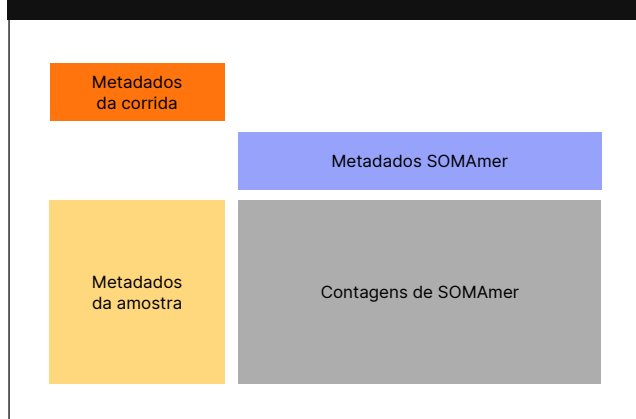
Figura 4: Impacto da normalização na variabilidade dos dados.



A normalização dos dados brutos do ensaio resulta em redução do coeficiente de variação (CV) intraplaca. Os dados foram gerados usando o ensaio Illumina Protein Prep 6K com amostras de CQ de plasma e soro. O ensaio Illumina Protein Prep 9.5K com CV mediano ~5,5%^a está agora disponível.

a. CV mediano esperado calculado usando amostras de doadores saudáveis.

Figura 5: Estrutura do arquivo ADAT.



Os arquivos ADAT são arquivos delimitados por tabulação com a seguinte estrutura (Figura 5):

- Os **metadados da corrida** contêm métricas que são relevantes para toda a corrida ou para placas individuais (por exemplo, RunID e métricas associadas à porcentagem de verificação de CQ nos extremos)
- Os **metadados do SOMAmer** contêm métricas que são relevantes para cada SOMAmer Reagent ou alvo de proteína (por exemplo, ID do SOMAmer e IDs específicas do gene e alvo)
- Os **metadados da amostra** contêm métricas que são relevantes para cada amostra (por exemplo, ID da amostra, ID da placa e Métricas de CQ)
- As **contagens de SOMAmer** são uma matriz de contagens de NGS que refletem a abundância relativa de cada SOMAmer em cada amostra

O formato foi projetado para fornecer flexibilidade para o número de amostras, bem como para o número e os tipos de descritores de analitos e amostras, e pode ser usado para análise terciária posterior para obter insights biológicos.

Relatório de CQ do DRAGEN

Um relatório de CQ do DRAGEN Protein Quantification acompanha o arquivo ADAT e resume o CQ de cada amostra individual e de toda a placa. O CQ de amostras individuais é derivado das etapas de normalização de hibridização e normalização mediana de referência externa. O CQ da placa é derivado das etapas de normalização da calibração e CQ.

Uso de arquivos ADAT para análise terciária

O arquivo ADAT normalizado pode ser usado imediatamente para análise adicional de dados com o Illumina Connected Multiomics via Illumina Connected Analytics. O Illumina Connected Multiomics é equipado com uma interface gráfica do usuário para gerenciamento e seleção de dados e pode ser usado para interpretação avançada de dados, incluindo expressão diferencial de proteína, análise de via e análise de enriquecimento de conjunto de genes. O conteúdo do arquivo ADAT também pode ser analisado usando pacotes de código aberto fornecidos em R (SomadataIO) e Python (Canopy).

Resumo

A normalização de dados proteômicos brutos é essencial para minimizar o viés do ensaio e da amostra que pode surgir durante o fluxo de trabalho do Illumina Protein Prep. O pipeline do DRAGEN Protein Quantification usa várias etapas de normalização e calibração para reduzir a variação sistemática, minimizando o risco de remover a verdadeira variação biológica, fornecendo resultados mais precisos. Esse pipeline está disponível em servidores na nuvem ou no local. Na nuvem, o pipeline do DRAGEN Protein Quantification é iniciado automaticamente após o sequenciamento nos sistemas NovaSeq 6000 ou NovaSeq X, gerando um relatório de CQ do DRAGEN e arquivos ADAT com contagens de proteínas normalizadas. A saída do arquivo ADAT pode, então, ser prontamente usada para análise terciária posterior com o Illumina Connected Multiomics para permitir insights biológicos mais profundos sobre o proteoma.

Saiba mais →

[Illumina Protein Prep](#)

[DRAGEN secondary analysis](#)

[Illumina Connected Analytics](#)



+1 (800) 809-4566, ligação gratuita (EUA) | tel. +1 (858) 202-4566
techsupport@illumina.com | www.illumina.com

© 2025 Illumina, Inc. Todos os direitos reservados. Todas as marcas comerciais pertencem à Illumina, Inc. ou aos respectivos proprietários. Para obter informações específicas sobre marcas comerciais, consulte www.illumina.com/company/legal.html.
M-GL-03205 PTB v1.0