

Consiga una detección de proteínas precisa y fiable con la solución Illumina Protein Prep



Una guía para los pasos de CC y normalización de datos utilizando el proceso de análisis secundario de DRAGEN™ Protein Quantification

Resultados reproducibles

Consiga una detección de proteínas uniforme en muestras y placas

Análisis de inicio automático

Genere resultados de alta calidad con el análisis secundario de DRAGEN

Información detallada integrada

Analice fácilmente archivos de recuento de proteínas normalizados con Illumina Connected Multiomics

Introducción

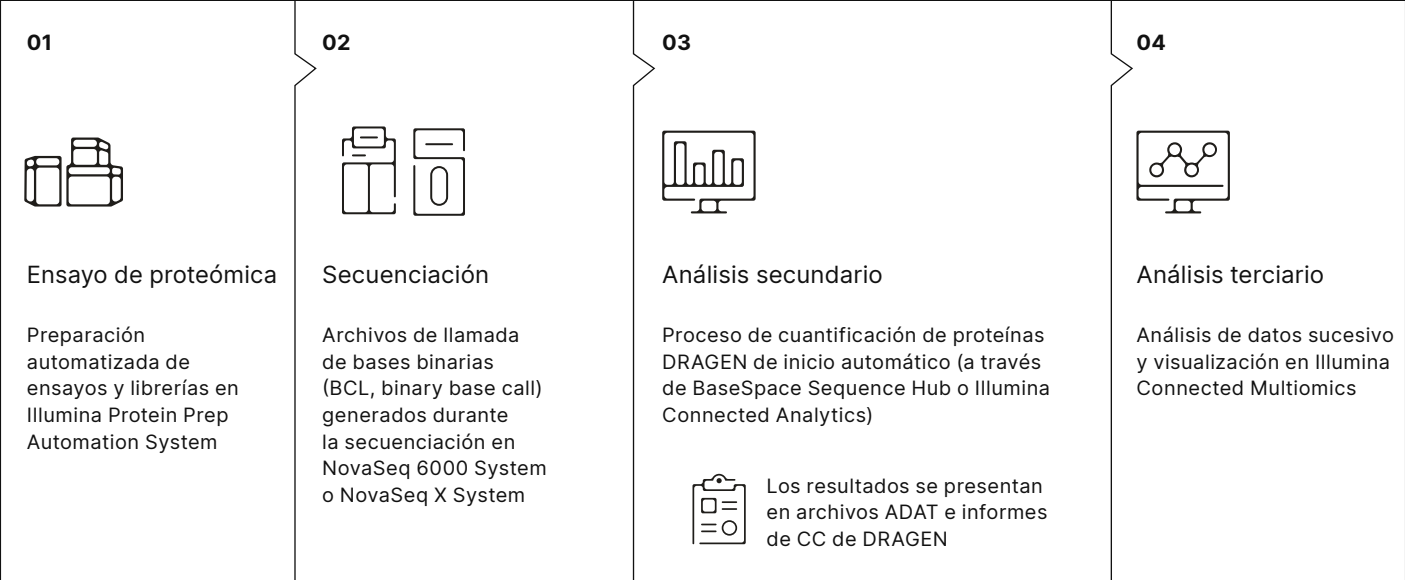
Las proteínas desempeñan un papel funcional clave en la biología humana y reflejan una instantánea en tiempo real de los estados de salud y enfermedad. La cuantificación precisa y reproducible de proteínas es crucial para obtener información biológica significativa. Illumina Protein Prep es un innovador ensayo de proteómica que utiliza reactivos SOMAmer® (aptámero modificado de disociación lenta) para capturar proteínas, lo que permite una detección de proteínas altamente sensible. La combinación de este ensayo con la lectura de secuenciación de nueva generación (NGS, next-generation sequencing) y la potencia bioinformática del análisis secundario de DRAGEN ofrece una solución proteómica optimizada y de alta productividad entre muestras y resultados.

Illumina Protein Prep detecta miles de proteínas simultáneamente en una sola muestra, lo que proporciona una potente plataforma para el descubrimiento de biomarcadores. Sin embargo, el sesgo entre placas y la variabilidad técnica de las plataformas proteómicas de alta productividad

pueden ocultar la expresión diferencial, especialmente para biomarcadores con tamaños de efecto pequeños. El proceso de DRAGEN Protein Quantification aborda este reto durante el análisis secundario, con métodos de normalización sólidos para ofrecer datos comparables de alta calidad en todos los experimentos. En la nube, el proceso se inicia automáticamente una vez finalizada la secuenciación en NovaSeq™ 6000 o NovaSeq X System, y genera un informe de control de calidad DRAGEN y archivos ADAT con recuentos de proteínas normalizados que se pueden utilizar para el análisis terciario posterior con Illumina Connected Multiomics (figura 1) para obtener información biológica más detallada sobre el proteoma. El proceso DRAGEN Protein Quantification se puede instalar en un servidor local de DRAGEN para los usuarios que necesiten una opción de análisis local.

En esta nota técnica se explican los pasos de control de calidad (CC) y normalización incorporados en el proceso de cuantificación de proteínas DRAGEN para eliminar la variabilidad técnica y de muestras para los datos proteómicos de alta calidad.

Figura 1: Descripción general del flujo de trabajo de Illumina Protein Prep.



La solución de proteómica de alta productividad utiliza el ensayo de proteómica basada en reactivos SOMAmer para la detección sensible de proteínas a partir de muestras de suero o plasma, seguida de la preparación y secuenciación de librerías de Illumina en NovaSeq 6000 System o NovaSeq X System. El ensayo de proteómica y el flujo de trabajo de preparación de librerías se automatizan en Illumina Protein Prep Automation System, un Tecan Fluent 780 personalizado, para obtener resultados coherentes y reproducibles. El proceso DRAGEN Protein Quantification se inicia automáticamente después de la secuenciación en BaseSpace Sequence Hub o Illumina Connected Analytics, transformando los datos de proteómica en archivos BCL producidos por los sistemas de secuenciación de Illumina en recuentos de proteómica normalizados en un archivo ADAT con un informe de CC de DRAGEN adjunto. El archivo ADAT se puede integrar directamente en programas de análisis terciario, incluido Illumina Connected Multiomics, para una interpretación adicional de los datos.

¿Por qué es necesaria la normalización de los datos?

La variabilidad técnica es inherente a cualquier ensayo proteómico de alta productividad y puede surgir de múltiples fuentes, incluida la preparación de muestras, la heterogeneidad intrínseca de las muestras, la eficiencia de la hibridación, la eficiencia de la PCR y la secuenciación. Además de la variabilidad técnica dentro de una muestra, pueden producirse efectos de lote que afectan a todas las muestras de la misma placa, debido a diferencias en los efectos del operador, el instrumento o el entorno. Sin una normalización eficaz o una corrección de lotes, estas variaciones sistemáticas podrían ocultar diferencias biológicas verdaderas, lo que daría lugar a conclusiones imprecisas.

La normalización de los datos es esencial para corregir esta variabilidad, estandarizar los datos y permitir comparaciones significativas entre muestras, placas e incluso diferentes ejecuciones de un mismo ensayo. La aplicación de estrategias de normalización completas con DRAGEN Protein Quantification proporciona datos de cuantificación de proteínas fiables y reproducibles.

Controles de Illumina Protein Prep

El ensayo Illumina Protein Prep incluye controles de placas y controles de reactivos SOMAmer en cada muestra para mitigar la variabilidad del ensayo. Ambos tipos de controles se utilizan para generar recuentos de expresión de proteínas normalizados con el proceso DRAGEN Protein Quantification.

Controles de placa

Cada placa de 96 pocillos tiene capacidad para 85 muestras y 11 controles de placa (figura 2), incluidos los blancos, los controles de calibrador y las muestras de CC. Estos controles se incluyen con el ensayo Illumina Protein Prep y se analizan en cada placa junto con las muestras de plasma o suero para la normalización y el CC. Sirven como puntos de referencia para detectar y corregir desviaciones técnicas y proporcionan una base para armonizar muestras procesadas en diferentes placas.

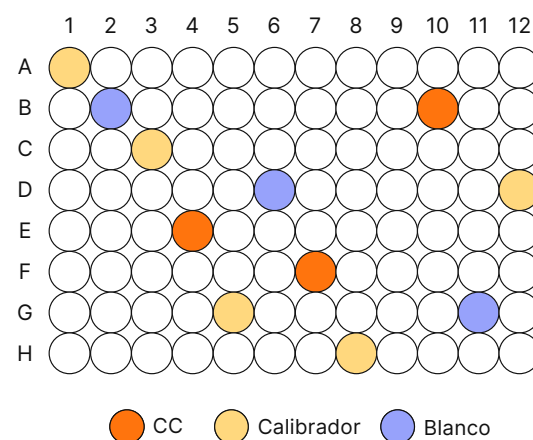
Blancos

Las muestras del blanco constan de tampón puro y proporcionan una estimación del fondo inespecífico en el ensayo Illumina Protein Prep. Se espera que estos controles tengan recuentos significativamente más bajos que las muestras de plasma o suero. Los recuentos elevados en los pocillos del blanco podrían indicar que un experimento tiene un fondo elevado o que se ha producido contaminación de la muestra.

Controles del calibrador

Los controles del calibrador permiten corregir la variabilidad entre placas. Estos controles son plasma o suero humano agrupado, en función de los tipos de muestras que se estén procesando. La normalización de la calibración de placas se realiza comparando la mediana de los controles del calibrador por reactivo SOMAmer con el control de calibración correspondiente.

Figura 2: Ejemplo de disposición de la placa con controles.



Muestras de CC

Las muestras de CC proporcionan una comprobación de control de calidad para determinar si la placa ha funcionado según lo previsto. Estas muestras de control son plasma o suero humano agrupado, en función de los tipos de muestras que se estén procesando. La calidad de la placa se evalúa comparando la mediana de estas muestras por reactivo SOMAmer con la referencia de CC correspondiente.

Controles de reactivos SOMAmer

Los controles de reactivos SOMAmer se incorporan en cada muestra individual a concentraciones conocidas antes de la hibridación y permiten calcular los factores de adaptación específicos de la muestra para normalizar las muestras y evaluar la calidad.

Pasos de normalización de datos

La normalización de los recuentos proteómicos sin procesar implica varios pasos para eliminar sesgos sistemáticos en las muestras y placas que pueden surgir durante el flujo de trabajo de Illumina Protein Prep. Estos ajustes se realizan en cuatro pasos no consecutivos: normalización de la hibridación, normalización de la mediana del control externo, calibración y adaptación de la placa (figura 3).

Normalización de la hibridación

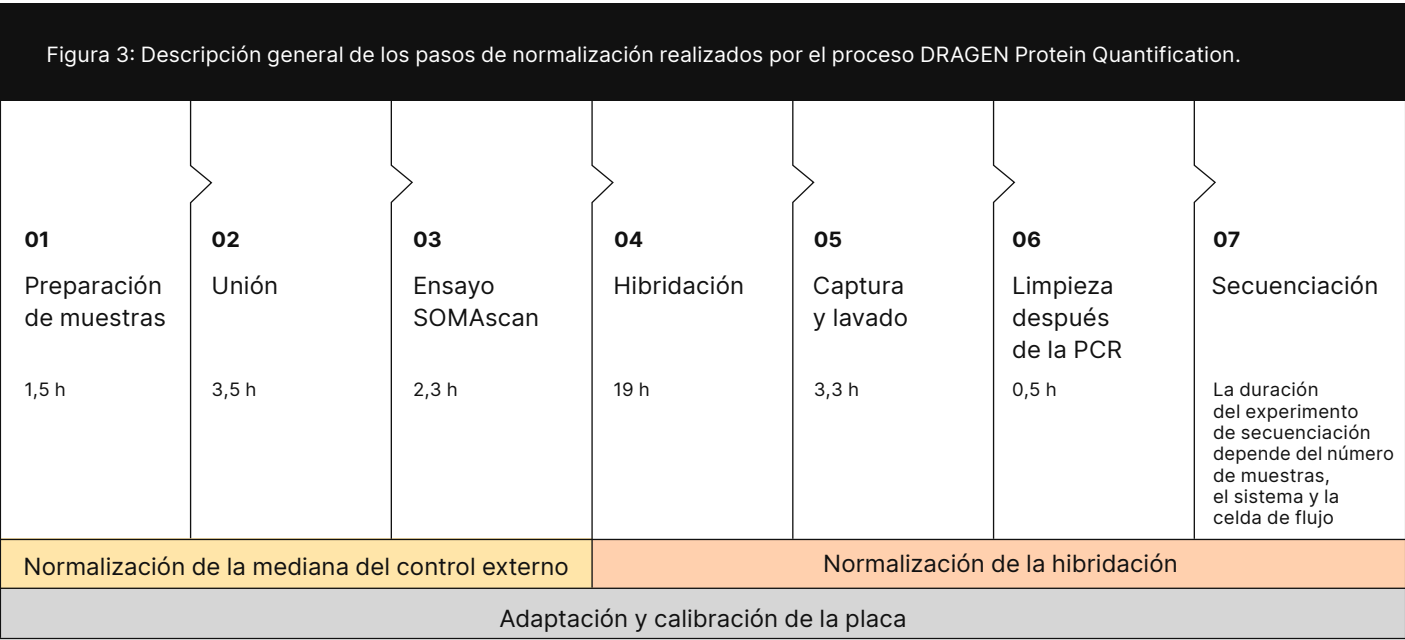
Este paso corrige la variabilidad entre muestras que puede producirse durante las etapas de hibridación y preparación de librerías del ensayo. Los factores de adaptación específicos de la muestra se calculan comparando los recuentos de los controles de normalización de la hibridación en cada muestra con la mediana de sus valores en todas las muestras de la placa.

Normalización de la mediana del control externo

Este paso corrige las diferencias en la medición de la abundancia total de proteínas de las muestras que no son controles. Un factor de adaptación se calcula comparando las mediciones de proteínas observadas con un control de los valores esperados para cada proteína.

Calibración

La calibración se realiza utilizando las cinco muestras de calibrador descritas en la sección de controles de la placa y corrige los efectos del lote que afectan a los reactivos SOMAmer individuales. Este paso permite realizar comparaciones precisas entre los datos producidos a partir de diferentes placas y procesamientos. Para cada reactivo SOMAmer, la calibración corrige la variación entre los calibradores y los valores de referencia esperados para el control del calibrador.



La calibración se lleva a cabo en dos pasos:

- Paso 1: se comparan las medianas del calibrador con un control obtenido del sistema de secuenciación y se ajusta cada reactivo SOMAmer en consecuencia.
- Paso 2: se comparan las medianas del calibrador actualizadas con un control derivado de los sistemas de secuenciación, lo que permite realizar comparaciones entre los datos generados en diferentes instrumentos. El factor de adaptación utilizado para armonizar la mediana de los calibradores con el valor de referencia se aplica a todas las muestras de la placa.

Adaptación de placas

La adaptación de placas aborda las variaciones en las mediciones de proteínas totales entre placas mediante el uso de la mediana de cada medición de reactivos SOMAmer entre calibradores y armonizándolas con un control de calibración externo.

El proceso de adaptación se lleva a cabo en dos pasos:

- Paso 1: se ajustan las mediciones en función de un control obtenido del instrumento de secuenciación, es decir, NovaSeq 6000 System o NovaSeq X System.
- Paso 2: se perfecciona el ajuste utilizando un control que se mantiene entre todos los sistemas de secuenciación, lo que permite comparaciones entre los datos generados en diferentes instrumentos. En el informe de DRAGEN Protein Quantification solo se muestran los criterios de medición del CC del primer paso de adaptación.

Porcentaje de comprobación de CC en las colas

Al final de la normalización se realiza un paso de CC con los controles de CC de cada placa. En este paso se compara la mediana de cada medición del reactivo SOMAmer en las tres réplicas de muestras de CC con un control de CC externo, a fin de asignar un factor de adaptación específico del reactivo SOMAmer y un criterio de medición de CC (QCCheckTailPercent).

Los diversos procedimientos de normalización implementados en el proceso DRAGEN Protein Quantification reducen al mínimo la variabilidad externa que se origina en distintas etapas del flujo de trabajo de Illumina Protein Prep (figura 4).

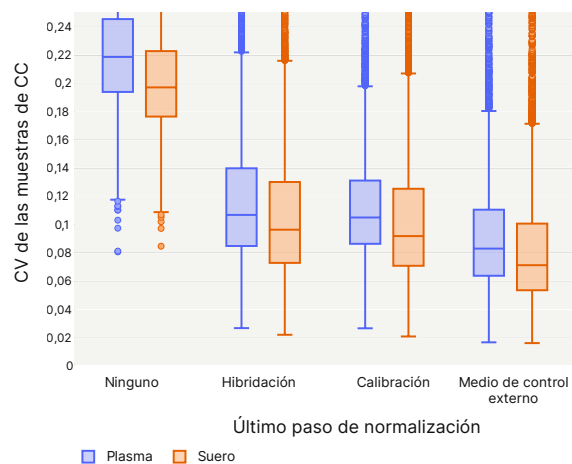
Rendimiento de DRAGEN Protein Quantification

El proceso DRAGEN Protein Quantification genera dos tipos de archivos que son fundamentales para analizar los datos de Illumina Protein Prep. El primero es una serie de archivos ADAT de cada paso de normalización, y el segundo es un informe de CC de DRAGEN que resume el CC del experimento.

Archivos ADAT

El proceso DRAGEN Protein Quantification procesa los recuentos sin procesar y genera un archivo de texto ASCII delimitado por tabulaciones con una extensión ADAT como resultado para cada paso de normalización realizado. El archivo ADAT contiene mediciones de una serie de analitos (columnas) en una serie de muestras (filas) e incluye la descripción del analito y una descripción de la muestra.

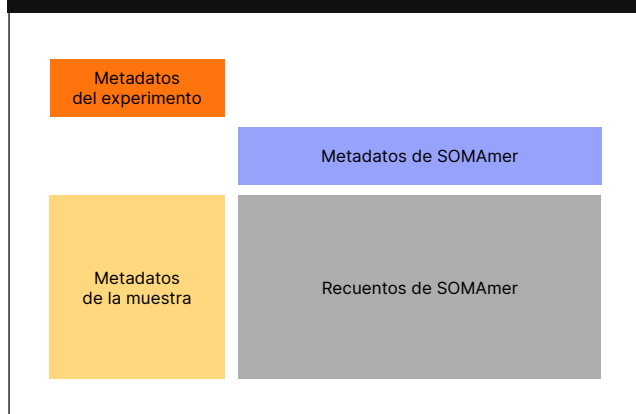
Figura 4: Impacto de la normalización en la variabilidad de los datos.



La normalización de los datos sin procesar del ensayo da como resultado una reducción del coeficiente de variación (CV) intraplaca. Los datos se generaron con el ensayo Illumina Protein Prep 6K con muestras de CC de plasma y suero. Ya está disponible el ensayo Illumina Protein Prep 9.5K con una mediana de CV aprox. 5,5 %^a.

a. Mediana prevista del CV calculado con muestras de donantes sanos.

Figura 5: Estructura de los archivos ADAT.



Los archivos ADAT son archivos delimitados por tabulaciones con la siguiente estructura (figura 5):

- **Los metadatos del experimento** incluyen criterios de medición relevantes para todo el experimento o las placas individuales (p. ej., ID del experimento y criterios de medición asociados al porcentaje de comprobación de CC en las colas).
- **Los metadatos de SOMAmer** incluyen criterios de medición relevantes para cada reactivo SOMAmer o proteína objetivo (p. ej., ID de SOMAmer y los ID específicos del gen y el objetivo).
- **Los metadatos de la muestra** incluyen criterios de medición relevantes para cada muestra (p. ej., ID de muestra, ID de placa y criterios de medición de CC).
- **Los recuentos de SOMAmer** son una matriz de recuentos de NGS que refleja la abundancia relativa de cada SOMAmer en cada muestra.

El formato está diseñado para proporcionar flexibilidad en cuanto al número de muestras, así como el número y los tipos de analito y descriptores de muestras, y se puede utilizar para el análisis terciario posterior para obtener información biológica detallada.

Informe de CC de DRAGEN

El archivo ADAT se acompaña de un informe de CC de DRAGEN Protein Quantification y resume el CC de cada muestra individual y de la placa completa. El CC de muestras individuales se obtiene de los pasos de normalización de la hibridación y normalización de la mediana de referencia externa. El CC de la placa se obtiene de los pasos de normalización de la calibración y el CC.

Uso de archivos ADAT para análisis terciarios

El archivo ADAT normalizado se puede utilizar de inmediato para realizar análisis de datos adicionales con Illumina Connected Multiomics a través de Illumina Connected Analytics. Illumina Connected Multiomics incluye con una interfaz gráfica de usuario para la gestión y selección de datos, y se puede utilizar para la interpretación avanzada de datos, incluida la expresión diferencial de proteínas, el análisis de rutas y el análisis de enriquecimiento del conjunto de genes. El contenido del archivo ADAT también se puede analizar utilizando paquetes de código abierto proporcionados en R (SomadataIO) y Python (Canopy).

Resumen

La normalización de los datos proteómicos sin procesar es esencial para reducir al mínimo el sesgo del ensayo y la muestra que puede surgir durante el flujo de trabajo de Illumina Protein Prep. El proceso DRAGEN Protein Quantification utiliza varios pasos de normalización y calibración para reducir la variación sistemática a la vez que minimiza el riesgo de eliminar la variación biológica real, lo que proporciona resultados más precisos. Este proceso está disponible en servidores en la nube o locales. En la nube, el proceso DRAGEN Protein Quantification se inicia automáticamente tras la secuenciación en NovaSeq 6000 System o NovaSeq X System, generando un informe de CC de DRAGEN y archivos ADAT con recuentos de proteínas normalizados. El resultado del archivo ADAT se puede utilizar fácilmente para el análisis terciario posterior con Illumina Connected Multiomics para obtener información biológica más detallada sobre el proteoma.

Más información →

[Illumina Protein Prep](#)

[Análisis secundario de DRAGEN](#)

[Illumina Connected Analytics](#)



1 800 809 4566 (llamada gratuita, EE. UU.) | tel.: +1 858 202 4566
techsupport@illumina.com | www.illumina.com

© 2025 Illumina, Inc. Todos los derechos reservados. Todas las marcas comerciales pertenecen a Illumina, Inc. o a sus respectivos propietarios. Si desea consultar información específica sobre las marcas comerciales, consulte www.illumina.com/company/legal.html.
M-GL-03205 ESP v1.0