

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 589 530**

51 Int. Cl.:

**G01N 33/68** (2006.01)

**C07K 19/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.04.2012 E 14157821 (1)**

97 Fecha y número de publicación de la concesión europea: **15.06.2016 EP 2767834**

54 Título: **Patrón cuantitativo para espectrometría de masas de proteínas**

30 Prioridad:

**04.04.2011 EP 11002794**

**04.04.2011 US 201161471528 P**

**04.04.2011 US 201161471534 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**15.11.2016**

73 Titular/es:

**MAX-PLANCK-GESELLSCHAFT ZUR  
FÖRDERUNG DER WISSENSCHAFTEN E.V.  
(50.0%)**

**Hofgartenstrasse 8  
80539 München, DE y  
ATLAS ANTIBODIES AB (50.0%)**

72 Inventor/es:

**MANN, MATTHIAS;  
STRAUBE, WERNER;  
ZEILER, MARLIS;  
UHLÉN, MATHIAS y  
LUNDBERG, EMMA**

74 Agente/Representante:

**ARIAS SANZ, Juan**

**ES 2 589 530 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Patrón cuantitativo para espectrometría de masas de proteínas

5 La presente invención se refiere a un método para determinar la cantidad absoluta de un péptido diana en una muestra, comprendiendo dicho método las siguientes etapas: (a) añadir (aa) un polipéptido de fusión a dicha muestra, comprendiendo dicho polipéptido de fusión (i) al menos una secuencia marcadora y (ii) una subsecuencia del polipéptido diana; y (ab) una cantidad absoluta conocida de un polipéptido marcador que comprende o consiste en dicha secuencia marcadora de acuerdo con (aa) a dicha muestra, en el que dicho polipéptido de fusión por un  
10 lado tiene masa alterada en comparación con dicho polipéptido diana y dicho polipéptido marcador por otro lado, por ejemplo, dicho polipéptido de fusión y dicho polipéptido diana y dicho polipéptido marcador por otro lado se marcan con isótopos de forma diferente; (b) realizar una digestión proteolítica de la mezcla obtenida en la etapa (a); (c) someter el resultado de la digestión proteolítica de la etapa (b), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y (d) determinar la cantidad absoluta de dicho polipéptido diana de (i) las intensidades  
15 pico en el espectro de masas adquirido en la etapa (c) de dicho polipéptido de fusión, dicho polipéptido marcador y dicho polipéptido diana y (ii) dicha cantidad absoluta conocida de dicho polipéptido marcador.

La proteómica basada en espectrometría de masas (EM) se ha convertido en un método elegido para estudiar proteínas de manera global (1-3). La espectrometría de masas no es inherentemente cuantitativa pero se han desarrollado métodos que aborden esta limitación en cierto grado. La mayoría de ellos se basan en isótopos estables e introducen una versión con masa desplazada de los péptidos de interés, que después se cuantifican por su relación de “pesado” frente a “ligero”. Se consigue marcaje isotópico estable mediante la adición química de reactivos marcados, marcaje con isótopos enzimáticos o marcaje metabólico (4-6). En general, estos enfoques se usan para obtener información cuantitativa relativa en los niveles de expresión de proteoma en una muestra  
20 marcada ligera y una pesada. Por ejemplo, el marcaje isotópico estable por aminoácidos en cultivo celular SILAC (7, 8) se realiza mediante incorporación metabólica de aminoácidos marcados de forma diferente, tal como marcados ligeros o pesados en el proteoma. Los proteomas marcados también pueden usarse como patrones internos para determinar los niveles de proteínas de un proteoma celular o tisular de interés, tal como en el enfoque de SILAC con adiciones (9).

30 La cuantificación absoluta es técnicamente más difícil que la cuantificación relativa y hasta la fecha solo ha podido realizarse con precisión para proteínas individuales o un número pequeño de proteínas cada vez (10). Son aplicaciones típicas de cuantificación absoluta la determinación de números de copias celulares de proteínas (importante para biología de sistemas) o la concentración de biomarcadores en fluidos corporales (importante para aplicaciones médicas). Además, cualquier método preciso de cuantificación absoluta, cuando se realiza en más de una muestra, también produce las cantidades relativas de la proteína entre estas muestras.

Han surgido varios métodos para la cuantificación absoluta durante los últimos años incluyendo AQUA (11), QConCAT (12, 13), PSAQ (14), SILAC absoluto (15) y FlexiQuant (16). Todos cuantifican la proteína endógena de interés por las relaciones de pesado con respecto a ligero hasta una cantidad definida del homólogo marcado  
40 añadido a la muestra y se distinguen principalmente entre sí añadiendo péptidos marcados pesados o proteínas de longitud completa marcadas pesadas. La estrategia AQUA usa péptidos prototípicos (17) que se sintetizan químicamente con isótopos pesados y se añaden después de la preparación de muestras. Los péptidos AQUA están disponibles en el mercado pero son caros, especialmente cuando es necesario cuantificar muchos péptidos o proteínas (véase, por ejemplo, Kettenbach *et al.*, Nat Protoc. 2011, 6: 175-86). Además, la estrategia AQUA padece incertidumbres de cuantificación que se introducen debido a la adición del patrón peptídico después de la preparación de muestra y proteólisis enzimática, que está en un estadio tardío en el flujo de trabajo. Además, cualquier pérdida de los péptidos, por ejemplo durante el almacenamiento, influiría directamente en los resultados de cuantificación. El enfoque de QconCAT se basa en proteínas artificiales que son concatémicos de péptidos proteotípicos. Esta proteína artificial se expresa de forma recombinante en *Escherichia coli* y se añade a la muestra antes de la proteólisis. QconCAT permite la producción de péptidos marcados, pero no corrige ningún desvío que surja de efectos de fraccionamiento de proteínas o eficacia de digestión. Los enfoques de PSAQ, SILAC absoluto y FlexiQuant intentan abordar estas limitaciones marcando metabólicamente proteínas de longitud completa por versiones pesadas de los aminoácidos arginina y lisina. PSAQ y FlexiQuant sintetizan proteínas de longitud completa *in vitro* en extractos de germen de trigo o en extracto de células bacterianas, respectivamente, mientras que SILAC absoluto se describió con expresión de proteína recombinante en *E. coli*. El patrón proteico se añade en un estadio temprano, tal como directamente al lisado celular. En consecuencia, puede realizarse fraccionamiento de muestra en paralelo y la proteína SILAC se digiere junto con el proteoma que se investiga. Sin embargo estas ventajas tienen el coste de tener que producir proteínas de longitud completa, lo que limita el rendimiento y en general restringe estos métodos a proteínas solubles.

En consecuencia, existe una necesidad no cumplida de medios mejorados o alternativos y métodos de cuantificación absoluta basada en espectrometría de masas de péptidos y polipéptidos.

65 Se describe un método para determinar la cantidad absoluta de un polipéptido diana en una muestra, comprendiendo dicho método las siguientes etapas: (a) añadir (aa) un polipéptido de fusión a dicha muestra,

comprendiendo dicho polipéptido de fusión (i) al menos una secuencia marcadora y (ii) una subsecuencia del polipéptido diana; y (ab) una cantidad absoluta conocida de un polipéptido marcador que comprende o consiste en dicha secuencia marcadora de acuerdo con (aa) a dicha muestra, en el que dicho polipéptido de fusión por un lado tiene la masa alterada en comparación con dicho polipéptido diana y dicho polipéptido marcador por otro lado, por ejemplo, dicho polipéptido de fusión por un lado y dicho polipéptido diana y dicho polipéptido marcador por otro lado se marcan con isotópicos diferentes; (b) realizar digestión proteolítica de la mezcla obtenida en la etapa (a); (c) someter el resultado de digestión proteolítica de la etapa (b), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y (d) determinar la cantidad absoluta de dicho polipéptido diana de (i) las intensidades pico en el espectro de masas adquirido en la etapa (c) de dicho polipéptido de fusión, dicho polipéptido marcador y dicho polipéptido diana y (ii) dicha cantidad absoluta conocida de dicho polipéptido marcador.

La expresión "cantidad absoluta" tiene su significado habitual y debe considerarse distinta de cantidades relativas, es decir relaciones, como se determinan habitualmente en análisis de expresión, sea mediante perfil de expresión de ARNm o métodos proteómicos. En particular, se entiende que la expresión "cantidad absoluta" se refiere al número de copias o la cantidad de sustancia de una proteína o un polipéptido dado en, por ejemplo, una célula, o la cantidad en un volumen definido, o en una muestra tal como ng/ml de un fluido corporal tal como orina o plasma. En otras palabras, dicha cantidad absoluta puede expresarse con respecto a una concentración, una masa o cantidad de sustancia (en moles o número de moléculas).

El término "polipéptido" está bien establecido en la técnica y se refiere a un policondensado de aminoácidos, preferentemente de los 20 aminoácidos convencionales. Se entiende que el término "polipéptido" como se usa en el presente documento abarca también péptidos, en los que los péptidos tienen una longitud mínima de dos aminoácidos. Por otro lado el término "polipéptido" incluye proteínas, al menos en la medida en que dichas proteínas consisten en una única cadena. Las proteínas a su vez también pueden comprender más de una cadena polipeptídica.

Se entiende que los métodos son igualmente adecuados para determinar las cantidades absolutas de proteínas, también en la medida en que las proteínas comprenden más de una cadena polipeptídica. En dicho caso, y suponiendo que se conocen las relaciones molares de las cadenas polipeptídicas comprendidas en la proteína, puede ser suficiente determinar la cantidad absoluta de un polipéptido comprendido en la proteína de interés. Como alternativa, la cantidad absoluta de más de uno o todos los polipéptidos comprendidos en la proteína de interés pueden determinarse por los métodos.

Un "polipéptido de fusión" de acuerdo con la invención es un polipéptido que comprende al menos dos segmentos de diferente origen. Más específicamente, un polipéptido de fusión de acuerdo con la invención requiere la presencia de una secuencia de aminoácidos marcadora y una subsecuencia del polipéptido diana comprendida o que se sospecha que está comprendida en la muestra indicada. Se prevé deliberadamente que está presente más de una secuencia de aminoácidos marcadora. Este es el objeto de realizaciones preferidas analizadas adicionalmente posteriormente. Además, esto se ejemplifica en los ejemplos incluidos y se representa en la Figura 1. Se describen realizaciones preferidas de los polipéptidos de fusión adicionalmente posteriormente e incluyen marcadores de identificación de epítomos proteicos (PrEST). Se prefiere que dicha secuencia marcadora se elija de modo que la digestión proteolítica del proteoma diana por un lado y de la secuencia marcadora por otro lado produzcan dos conjuntos distintos de péptidos o al menos dos conjuntos de péptidos que se solapan en menos del 25 %, menos del 10 %, menos del 5 %, menos del 2 % o menos del 1 %. Un "proteoma diana" es típicamente un proteoma que se origina de una única especie. Un proteoma diana comprende dichos polipéptidos diana. Un proteoma preferido es un proteoma humano. Si está presente más de una secuencia marcadora, se entiende que las secuencias marcadoras son diferentes entre sí. En particular, el conjunto de péptidos obtenido por digestión proteolítica de una primera secuencia marcadora presente en dicho polipéptido de fusión y el conjunto de péptidos obtenido por digestión proteolítica de una segunda secuencia marcadora (y también cualquier secuencia marcadora adicional) presente en dicho polipéptido de fusión son distintos, es decir, no comparten un péptido de la misma secuencia. Siempre que se haga referencia en el presente documento a conjuntos distintos de péptidos obtenidos por digestión proteolítica, se entiende que los conjuntos de péptidos son distintos en particular con respecto a péptidos de o por encima de una longitud mínima, siendo dicha longitud mínima al menos 4, 5, 6, 7, 8 o 9 aminoácidos.

El término "subsecuencia" en su forma más amplia se refiere a cualquier secuencia parcial de un polipéptido diana para detectar y además incluye la secuencia completa de dicho polipéptido diana. En una realización preferida, dicha subsecuencia es una secuencia parcial del polipéptido diana, excluyéndose la secuencia completa de dicho polipéptido diana. Se analizan adicionalmente posteriormente intervalos de longitud preferidos de dicha subsecuencia.

El término "isótopo" se refiere a dos o más núclidos con el mismo número de protones (número atómico) pero diferente número de neutrones. Dicha diferencia en el número másico posibilita diferentes posiciones pico de un compuesto marcado con isótopo o fragmento por un lado y su homólogo no marcado por el otro en un espectro de masas. Son isótopos preferidos deuterio,  $^{13}\text{C}$  y  $^{15}\text{N}$ .

El término "marcado" se refiere a una frecuencia de isótopos que se desvía de la frecuencia de origen natural. En

realizaciones preferidas, la expresión “marcado con isótopo” se refiere a un compuesto, resto, fragmento o molécula que, en la medida en que se consideran átomos con el mismo número atómico, contiene exclusivamente un isótopo dado. Por ejemplo, una lisina marcada con isótopo preferida tiene núclidos de  $^{13}\text{C}$  en todas las posiciones de carbono. En realizaciones preferidas, uno o más aminoácidos específicos, tales como todas las lisinas y/o todas las argininas, están marcados con isótopos. Los restos de aminoácidos marcados con isótopos adecuados se enumeran adicionalmente posteriormente.

La expresión “marcado de forma diferente” o “marcado con isótopos de forma diferente” como se usa en el presente documento se refiere a una pluralidad de esquemas de marcaje. En particular, es suficiente para que dos polipéptidos estén marcados de forma diferente que uno de ellos esté marcado y el otro no. Se prevé igualmente que uno de los polipéptidos esté marcado con isótopo de una manera específica, mientras que el otro polipéptido está marcado con isótopo también, pero de una manera diferente, siendo la consecuencia que ambos polipéptidos no muestran la frecuencia de origen natural de isótopos y pueden distinguirse en el espectro de masas. Se entiende que “marcado con isótopos de forma diferente” de acuerdo con la invención es tal que, tras la digestión proteolítica, (i) se forma al menos un primer péptido del polipéptido diana y se forma al menos un segundo péptido de la subsecuencia del mismo como está comprendida en el polipéptido de fusión de modo que el primer y segundo péptido son de secuencia idéntica pero difieren en su masa, y (ii) se forma al menos un tercer péptido a partir del polipéptido marcador y se forma al menos un cuarto péptido a partir de la secuencia marcadora como está comprendida en el polipéptido de fusión de modo que el tercer y cuarto péptidos sean de secuencia idéntica pero difieran en su masa. Esto puede conseguirse, por ejemplo, mediante los polipéptidos marcados, que comprenden marcadores internos, preferentemente cada aparición de uno o más aminoácidos dados que se marcan, siendo dichos aminoácidos dados preferentemente los que están comprendidos en el sitio de escisión reconocido por la enzima usada para digestión proteolítica. Dichos aminoácidos preferidos son como se describe en otra parte en el presente documento, lisina y/o arginina. Tomados juntos, se prefiere que dicho polipéptido de fusión por un lado y dicho polipéptido diana y dicho polipéptido marcador por otro lado estén marcados con isótopos de forma interna diferente. Se entiende que el término “interno” como se usa en el presente documento en relación con marcadores se distingue de marcadores terminales.

En general, siempre que se haga referencia a “marcado de forma diferente” o a “marcado con isótopos de forma diferente” en la presente divulgación, se entiende que estas expresiones se refieren a una realización preferida. Más en general, se prevé cualquier medio de alteración de la masa incluyendo, pero sin limitación, marcaje isotópico. Las expresiones “alteración de la masa” y “de masa alterada” como se usan en el presente documento se refieren a todos los medios y métodos que proporcionan péptidos (o polipéptidos) obtenidos de diferentes fuentes y de secuencia idéntica que difieran con respecto a su masa. El marcaje isotópico es un medio preferido para conseguir este objetivo. Un método alternativo conocido en la técnica es el uso de marcadores isobáricos para cuantificación relativa y absoluta (iTRAQ). Este método usa marcadores covalentes codificados por isótopos; véase por ejemplo Ross *et al.*, Mol. Cell. Proteomics 3, 1154-69, 2004. Preferentemente, iTRAQ se basa en un marcaje covalente del extremo N terminal y aminas de cadenas laterales de péptidos y polipéptidos. Se conocen en la técnica agentes adecuados, cuyos ejemplos incluyen agentes denominados cuádruples y óctuples. Si se indica en el presente documento que una entidad A tiene masa alterada en comparación con una entidad B, se entiende que la entidad A o la entidad B se desvían de la forma de origen natural, por ejemplo por diferente marcaje de isótopo o debido a la presencia de marcadores covalentes en el sentido de iTRAQ.

Volviendo al requisito indicado anteriormente de que “al menos dicho polipéptido de fusión por un lado y dicho polipéptido diana y dicho polipéptido marcador por otro lado están marcados con isótopos de forma diferente”, se observa que dicho polipéptido diana y dicho polipéptido marcador puede marcarse con isótopos del mismo modo o de acuerdo con diferentes patrones de marcaje o, si dicho polipéptido de fusión está marcado con isótopos, ambos pueden estar no marcados. Más específicamente, se aceptan al menos los siguientes esquemas de marcaje. (1) Dicho polipéptido de fusión se marca con isótopo, y tanto dicho polipéptido diana como dicho polipéptido marcador no se marcan con isótopo, (2) dicho polipéptido diana y dicho polipéptido marcador se marcan con isótopo, y dicho polipéptido de fusión no se marca con isótopo, en el que el polipéptido diana y polipéptido marcador se marcan con isótopo de la misma manera o de acuerdo con diferentes patrones de marcaje, (3) un polipéptido seleccionado de polipéptido diana, polipéptido de fusión y polipéptido marcador no se marca con isótopo o se marca con isótopo de acuerdo con un primer patrón, un segundo polipéptido seleccionado del mismo grupo se marca con isótopo de acuerdo con un segundo patrón y el polipéptido restante del grupo se marca con isótopo de acuerdo con un tercer patrón. Los tres patrones (o dos patrones en el caso de que uno de los polipéptidos no esté marcado con isótopo) de acuerdo con el esquema de marcaje (3) pueden implementarse, por ejemplo, usando dos o tres formas marcadas con isótopo de uno o más aminoácidos dados, difiriendo dichas dos o tres formas marcadas con isótopo en la masa total. Un esquema de marcaje ejemplar de acuerdo con (3) es el siguiente: el polipéptido diana no está marcado con isótopo, el polipéptido de fusión está marcado con isótopo (forma de “peso pesado”) y el polipéptido marcador está marcado con un isótopo de acuerdo con un patrón diferente de modo que se proporciona, por ejemplo, en una forma de “peso medio” o una forma de “peso extrapesado”. Dicho esquema de marcaje puede preferirse particularmente si se sospecha que un producto proteolítico del polipéptido marcador también podría derivar de la digestión de la muestra, por ejemplo si la muestra es humana y el marcador es una proteína humana o un dominio o segmento de la misma.

La expresión “esquema de marcaje” como se usa en el presente documento distingue entre diferentes polipéptidos. Para un esquema de marcaje dado, una clase de polipéptidos (siendo las clases polipéptidos diana, polipéptidos marcadores y polipéptidos de fusión) se marca de la misma manera, por ejemplo por incorporación de una lisina marcada con  $^{13}\text{C}$  en todas las posiciones en las que aparece una lisina. Un esquema de marcaje posibilita marcar de forma diferente diferentes clases. Por otro lado, la expresión “patrón de marcaje” distingue entre formas marcadas de forma diferente de un péptido dado. Por ejemplo, un polipéptido específico puede marcarse reemplazando todas las apariciones de lisina con lisina marcada con  $^{13}\text{C}$  o reemplazando todas las posiciones de arginina con arginina marcada con  $^{13}\text{C}$   $^{15}\text{N}$ , haciendo de este modo a los patrones de marcaje diferentes.

Están disponibles para los expertos en la materia diversos medios de marcaje con isótopos e incluyen adición química de reactivos marcados, marcaje con isótopos enzimáticos o marcaje metabólico (4-6).

Se prefiere que el marcaje isotópico se introduzca mediante marcaje metabólico. En otras palabras, los polipéptidos para usar en los métodos de acuerdo con la invención, en la medida en que se requiere que se marquen, se obtienen preferentemente por medio de producción en sistemas biológicos, tal como sistemas sin células así como celulares. Por ejemplo, puede usarse una célula hospedadora que es autotrófica para lisina y/o arginina, en la que al mismo tiempo se proporciona lisina y/o arginina marcada con isótopo en el medio de cultivo. Un medio preferido de marcaje isotópico metabólico es marcaje isotópico estable con aminoácidos en cultivo celular (SILAC). Se conocen en la técnica procedimientos de SILAC y se describen en la sección de antecedentes del presente documento anterior así como en las referencias citadas en relación con la misma que se incorporan con la presente por referencia. Como se ha mencionado anteriormente, en la medida en que el marcaje isotópico hace uso de isótopos con mayores números másicos, la forma marcada se denomina habitualmente forma “pesada”, mientras que el homólogo de origen natural o el homólogo que está libre o esencialmente libre del isótopo pesado que se considera se denomina habitualmente forma “ligera”.

La indicada “cantidad absoluta conocida de un polipéptido marcador” puede determinarse con métodos establecidos en la técnica. Un método preferido es el análisis de aminoácidos. El análisis de aminoácidos se proporciona típicamente como un servicio por una diversidad de compañías. El método preferentemente incluye la hidrólisis total de una muestra dada, la derivatización química de los aminoácidos libres obtenidos, la separación de los aminoácidos derivatizados, por ejemplo mediante HPLC de fase inversa, y la interpretación posterior del resultado. El método se describe en más detalle, por ejemplo, en Moore y Stein, J. Biol. Chem. 176, 367-388 (1948) así como en Moore y Stein, J. Biol. Chem. 176, 337-365 (1948).

Los métodos requieren, por un lado, que una primera subsecuencia del polipéptido de fusión sea idéntico a una subsecuencia de la proteína diana, y por otro lado, que una segunda subsecuencia del polipéptido de fusión sea idéntica al polipéptido marcador. Además, incluso aunque las secuencias de aminoácidos son idénticas, es necesario que las masas de la primera subsecuencia del polipéptido de fusión y su homólogo en el polipéptido diana sean distintas. De forma similar, también es necesario que las masas de la segunda secuencia del polipéptido de fusión y el polipéptido marcador sean distintas. Esto puede conseguirse por los esquemas de marcaje descritos anteriormente. Esto permite realizar comparaciones cuantitativas entre la secuencia marcadora dentro del polipéptido de fusión y el polipéptido marcador así como entre dicha subsecuencia comprendida en dicho polipéptido de fusión y el polipéptido diana.

La etapa (b) proporciona digestión proteolítica que, como se ha establecido bien en la técnica, da lugar a fragmentos que pueden manipularse convenientemente en espectroscopia de masas. Se describen además posteriormente enzimas preferidas para usar para digestión proteolítica. Se prefiere que dicha digestión proteolítica sea específica, es decir, que se produzca escisión en todos los sitios de escisión de la enzima usada. Por otro lado, y como se describe en el presente documento, los métodos de la presente invención posibilitan la evitación de sesgo introducido por digestión incompleta.

Después de la digestión proteolítica, se realiza análisis de espectrometría de masas. Se transfieren moléculas peptídicas ionizadas a los sistemas de vacío del espectrómetro de masas. En un modo preferido de funcionamiento, ampliamente conocido para los practicantes de la técnica, el espectrómetro de masas se opera de modo que realice una exploración espectrométrica de masas que registre un espectro de masas de los péptidos que entran en el instrumento en ese momento. La cuantificación se basa en los picos presentes en esta exploración espectrométrica de masas (o EM). Los ejemplos incluidos proporcionan una explicación más detallada de modos adecuados de funcionamiento del espectrómetro de masas. Dependiendo de la naturaleza de las muestras para analizar, los polipéptidos que se sospecha que están comprendidos en la muestra y la instrumentación disponible, el experto en la materia puede elegir modos adecuados de funcionamiento.

Dado que se realiza digestión proteolítica, el polipéptido marcador que comprende dicha secuencia marcadora de acuerdo con (aa) o un polipéptido marcador que consiste en dicha secuencia marcadora de acuerdo con (aa) pueden usarse indistintamente. Preferentemente, en cualquiera de los casos se producirán el mismo o los mismos fragmentos marcadores durante la digestión proteolítica.

Antes de realizar análisis de espectrometría de masas, el resultado de digestión proteolítica puede someterse a

cromatografía como está establecido en la técnica. Son medios preferidos de cromatografía, cromatografía líquida (CL). En un modo preferido de funcionamiento, la mezcla de péptidos se inyecta en una columna de cromatografía líquida, separada por un gradiente de disolvente orgánico que dura varios minutos o varias horas y se electropulveriza en línea.

5 La etapa (d) combina la información obtenida en el espectro de masas (que puede verse como intensidades relativas) con la cantidad absoluta conocida del polipéptido marcador para determinar cantidad absolutas, en particular la cantidad absoluta del polipéptido diana comprendido en la muestra. Para explicar adicionalmente, y usar la terminología de primer a cuarto péptidos introducidos en el presente documento anteriormente, la cantidad absoluta de un polipéptido diana dado puede determinarse, por ejemplo, de la siguiente manera. Las relaciones de cantidades de sustancias son idénticas a relaciones de intensidades en el espectro de EM de los picos correspondientes. Usando los números de 1 a 4 como designaciones abreviadas de péptido primero a cuarto, se aplica lo siguiente. La cantidad de sustancia del cuarto péptido (fragmento proteolítico derivado de la secuencia marcadora comprendida en el polipéptido de fusión)  $N(4)$  puede determinarse de acuerdo con  $N(4) = N(3) \text{ por } I(4) / I(3)$ .  $N(3)$  es la cantidad absoluta conocida del polipéptido marcador.  $I(3)$  e  $I(4)$  son las intensidades pico correspondientes. Dada la definición del polipéptido de fusión, se aplica  $N(2) = N(4)$ , es decir, las cantidades de sustancia de los péptidos formados a partir de una de las partes del polipéptido de fusión son idénticas. La cantidad de sustancia del polipéptido diana  $N(1)$  puede determinarse de la siguiente manera:  $N(1) = N(2) \text{ por } I(1) / I(2)$ . Haciendo uso de  $N(2) = N(4)$  y  $N(4) = N(3) \text{ por } I(4) / I(3)$ , se concluye que  $N(1) = N(3) [I(1) I(4) / I(2) I(3)]$  que permite la cuantificación absoluta de los polipéptidos diana basándose en intensidades pico  $I(1)$  a  $I(4)$  y la cantidad absoluta conocida del polipéptido marcador  $N(3)$ . Obsérvese que en la práctica las relaciones se determinan habitualmente como la media de las relaciones de varias intensidades peptídicas; es decir más de un par de péptidos que abarcan la secuencia marcadora y la secuencia polipeptídica diana.

25 Los métodos hacen uso de esquemas de marcaje específicos de tres especies distintas, habiéndose descrito anteriormente los esquemas de marcaje. Una característica clave de los métodos de la invención es el uso de polipéptidos de fusión, conteniendo dichos polipéptidos de fusión al menos una secuencia genérica, también denominada "secuencia marcadora" en el presente documento. La condición conjunta de un polipéptido marcador como se ha definido anteriormente en una cantidad absoluta conocida permite la calibración de una manera que provechosamente es independiente del polipéptido real para determinar cuantitativamente.

35 Desviándose de una diversidad de métodos de la técnica anterior como se ha analizado anteriormente, los métodos de la presente invención posibilitan la adición temprana del patrón (en el caso de la realización principal de dicha cantidad absoluta conocida de un polipéptido marcador) en el flujo de trabajo completo. Como consecuencia, se aplican igualmente etapas corriente abajo que incluyen digestión proteolítica y opcionalmente cromatografía tanto al patrón como a los constituyentes de la muestra para analizar. Cualquier variación en la eficacia o rendimiento de, por ejemplo, la digestión proteolítica, afectará igualmente a todos los constituyentes de la mezcla obtenida en la etapa (a), evitando de este modo cualquier sesgo que pueda surgir de la misma. Preferentemente, no se usa ningún método basado en los tamaños de las proteínas tal como cromatografía de exclusión por tamaño después de dicha adición.

45 Se conocen bien por los practicantes de la proteómica que la cuantificación precisa de proteínas de muy baja abundancia es difícil. Sin embargo, la precisión de la cuantificación del patrón proteico de fusión en sí mismo no depende de la abundancia celular u otros atributos del polipéptido para determinar, observando que la misma cantidad de polipéptido de fusión se usa preferentemente en cada caso de los métodos de acuerdo con la invención. Además, la pureza de una composición que comprende dicho polipéptido de fusión para añadir no tiene ninguna influencia debido a que los métodos determinan específicamente la cantidad de polipéptido de fusión y no de proteína total.

50 Como se analiza en más detalle en los ejemplos incluidos con la presente, los métodos proporcionan precisión significativamente mejorada en la determinación cuantitativa de los niveles de expresión de proteína celular. Son ventajas adicionales del método que típicamente da como resultado varios péptidos cuantificables para cada polipéptido de fusión, tanto para la cuantificación precisa del patrón como para cuantificar de forma absoluta el polipéptido diana. Además, la producción de patrón puede simplificarse porque la expresión de proteínas puede realizarse en un sistema convencional (tal como *E. coli*) y debido a que puede producirse un gran número de polipéptido de fusión en condiciones similares ya que solamente difieren en una única secuencia relativamente corta en la realización preferida.

60 También se describe un método para crear un patrón cuantitativo, comprendiendo dicho método las siguientes etapas: (a) proporcionar una pluralidad de polipéptidos de fusión, comprendiendo cada uno de dichos polipéptidos de fusión (i) al menos una secuencia marcadora y (ii) una subsecuencia de un polipéptido diana para determinar de forma cuantitativa, en el que todos los polipéptidos de fusión comparten al menos una secuencia marcadora, obteniendo de este modo el patrón; (b) determinar las cantidades absolutas de dichos polipéptidos de fusión mediante (ba) la adición a uno de dichos polipéptidos de fusión a la vez una cantidad conocida de un polipéptido marcador que comprende o consiste en la secuencia marcadora compartida entre los polipéptidos de fusión de acuerdo con (a), en el que dicho polipéptido de fusión tiene masa alterada en comparación con dicho polipéptido

marcador, por ejemplo, dicho polipéptido de fusión y dicho polipéptido marcador están marcados con isótopo de forma diferente, (bb) realizar digestión proteolítica de la mezcla de un polipéptido de fusión y dicho polipéptido marcador obtenido en la etapa (ba); (bc) someter el resultado de digestión proteolítica de la etapa (bb), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y (bd) determinar la cantidad absoluta de dicho polipéptido de fusión de (i) las intensidades pico en el espectro de masas de polipéptido de fusión y polipéptido marcador y (ii) dicha cantidad conocida de dicho polipéptido marcador, obteniendo de este modo la cantidad absoluta de uno de dichos polipéptidos de fusión a la vez.

Aunque este último método posibilita la opción de multiplexar como se analiza adicionalmente posteriormente, debe observarse que dicho segundo aspecto no se restringe al uso de una pluralidad de polipéptidos de fusión. En consecuencia, también se describe un método para crear un patrón cuantitativo, comprendiendo dicho método las siguientes etapas: (a) proporcionar un polipéptido de fusión, comprendiendo el polipéptido de fusión (i) al menos una secuencia marcadora y (ii) una subsecuencia de un polipéptido diana para determinar de forma cuantitativa, obteniendo de este modo el patrón; (b) determinar la cantidad absoluta de dicho polipéptido de fusión mediante (ba) la adición al polipéptido de fusión de una cantidad conocida de un polipéptido marcador que comprende o consiste en la secuencia marcadora comprendida en el polipéptido de fusión de acuerdo con (a) en el que dicho polipéptido de fusión tiene masa alterada en comparación con dicho polipéptido marcador, por ejemplo, dicho polipéptido de fusión y dicho polipéptido marcador están marcados con isótopos de forma diferente, (bb) realizar digestión proteolítica de la mezcla de un polipéptido de fusión y dicho polipéptido marcador obtenido en la etapa (ba); (bc) someter el resultado de digestión proteolítica de la etapa (bb), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y (bd) determinar la cantidad absoluta de dicho polipéptido de fusión de (i) las intensidades pico en el espectro de masas de polipéptido de fusión y polipéptido marcador y (ii) dicha cantidad conocida de dicho polipéptido marcador, obteniendo de este modo la cantidad absoluta del polipéptido de fusión.

En otras palabras, parte de una preparación de polipéptidos de fusión se combina con una cantidad conocida de un polipéptido marcador, en el que el polipéptido de fusión tiene masa alterada en comparación con el polipéptido marcador. Esta mezcla binaria se somete a digestión proteolítica, análisis espectrométrico de masas y cuantificación para proporcionar la cantidad absoluta de la parte de polipéptidos de fusión, a partir de la que puede calcularse la concentración exacta del polipéptido de fusión en la preparación. Por lo tanto, se ha proporcionado un patrón cuantitativo de un único polipéptido de fusión. Entonces se añade al menos parte del patrón cuantitativo a la muestra para analizar, después de lo cual se realiza digestión proteolítica de la mezcla obtenida. El resultado de digestión proteolítica se somete a análisis espectrométrico de masas, opcionalmente después de cromatografía. La cantidad absoluta del polipéptido diana se determina después a partir de (i) las intensidades pico en el espectro de masas del polipéptido de fusión y el polipéptido diana y (ii) las cantidades absolutas conocidas del polipéptido de fusión, en el que dicho polipéptido de fusión tiene masa alterada en comparación con dicho polipéptido diana.

Por lo tanto, se entiende que este último método, que abarca de una forma más concisa tanto el uso de un polipéptido de fusión como una pluralidad de los mismos, se refiere a un método para crear un patrón cuantitativo, comprendiendo dicho método las siguientes etapas: (a) proporcionar uno o una pluralidad de polipéptidos de fusión, comprendiendo el polipéptido de fusión o cada uno de dichos polipéptidos de fusión, respectivamente, (i) al menos una secuencia marcadora y (ii) una subsecuencia de un polipéptido diana para determinar de forma cuantitativa, en el que, en la medida en que se proporciona dicha pluralidad de polipéptidos de fusión, todos los polipéptidos de fusión comparten al menos una secuencia marcadora, obteniendo de este modo el patrón; (b) determinar las cantidades absolutas de dichos polipéptidos de fusión (ba) añadiendo al polipéptido de fusión o a uno de dichos polipéptidos de fusión a la vez, respectivamente, de una cantidad conocida de un polipéptido marcador que comprende o consiste en la secuencia marcadora compartida en el polipéptido de fusión o compartida entre los polipéptidos de fusión, respectivamente, de acuerdo con (a), en el que dicho polipéptido de fusión tiene masa alterada en comparación con dicho polipéptido marcador, por ejemplo, dicho polipéptido de fusión y dicho polipéptido marcador están marcados con isótopo de forma diferente, (bb) realizando digestión proteolítica de la mezcla de un polipéptido de fusión y dicho polipéptido marcador obtenido en la etapa (ba); (bc) sometiendo el resultado de digestión proteolítica de la etapa (bb), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y (bd) determinando la cantidad absoluta de dicho polipéptido de fusión de (i) las intensidades pico en el espectro de masas de polipéptido de fusión y polipéptido marcador y (ii) dicha cantidad conocida de dicho polipéptido marcador, obteniendo de este modo la cantidad absoluta del polipéptido de fusión o de uno de dicha pluralidad de polipéptidos de fusión a la vez, respectivamente.

En relación con lo mismo, se describe también un método para determinar la cantidad absoluta de uno o más polipéptidos diana en una muestra, comprendiendo dicho método las siguientes etapas: (a) opcionalmente realizar el método de acuerdo con el segundo aspecto; (b) añadir el patrón cuantitativo como se ha definido en el segundo aspecto a dicha muestra; (c) realizar digestión proteolítica de la mezcla obtenida en la etapa (b); (d) someter el resultado de digestión proteolítica de la etapa (c), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y (e) determinar las cantidades absolutas del polipéptido o los polipéptidos diana de (i) las intensidades pico en el espectro de masas adquirido en la etapa (d) de polipéptido o polipéptidos de fusión y polipéptidos diana y (ii) la cantidad o las cantidades absolutas conocidas de dicho polipéptido o dichos polipéptidos de fusión, en el que dicho polipéptido o dichos polipéptidos de fusión tienen la masa alterada en comparación con dicho polipéptido o dichos polipéptidos diana, por ejemplo, dichos uno o más polipéptidos diana están marcados con

isótopos de forma diferente en comparación con dicho polipéptido o dichos polipéptidos de fusión.

Aunque el método descrito para determinar la cantidad absoluta de un polipéptido diana proporciona cuantificación absoluta de un polipéptido a partir de un único experimento de masa, los métodos para crear un patrón cuantitativo y para determinar la cantidad absoluta de uno o más polipéptidos diana, respectivamente, se refieren a (i) preparación y cuantificación de un patrón y (ii) uso de este patrón en la cuantificación de uno o más de una pluralidad de polipéptidos comprendidos en una muestra. Resulta importante que dicho enfoque es susceptible de multiplexación. En otras palabras, no solamente uno, sino también una pluralidad de polipéptidos comprendidos en una muestra puede determinarse conjuntamente de manera cuantitativa.

De acuerdo con el método para crear un patrón cuantitativo, se proporciona uno o una pluralidad de polipéptidos de fusión. De acuerdo con la etapa (b) del mismo, un polipéptido de fusión en el momento se combina con una cantidad conocida de un polipéptido marcador. Esta mezcla binaria se somete a digestión proteolítica, análisis espectrométrico de masas y cuantificación para proporcionar la cantidad absoluta de uno de dichos polipéptidos de fusión a la vez. Realizando la etapa (b) para el, los o todos los polipéptidos de fusión comprendidos en el patrón, el patrón se caracteriza cuantitativamente y puede usarse en un método de acuerdo con el tercer aspecto de la presente invención. El método del segundo aspecto proporciona en la etapa (a) la fabricación física del patrón cuantitativo y en la etapa (b) su caracterización con respecto a cantidades absolutas del polipéptido o los polipéptidos de fusión constituyentes. Los patrones cuantitativos preferidos también se denominan "mezcla maestra PrEST" en el presente documento.

Un método para determinar la cantidad absoluta de uno o más polipéptidos diana puede, de acuerdo con la etapa (a), incorporar el método para crear un patrón cuantitativo en su totalidad. Como alternativa, la etapa (a) puede omitirse. En ese caso, se entiende que el patrón cuantitativo para añadir de acuerdo con la etapa (b) se caracteriza de acuerdo con la etapa (b) del método para crear un patrón cuantitativo.

En consecuencia, el patrón interno (es decir el polipéptido de fusión) se cuantifica de este modo en una primera etapa usando un patrón interno del patrón interno (es decir el polipéptido marcador), y una proteína diana en una muestra se cuantifica en una segunda etapa posterior usando el patrón interno cuantificado (es decir el polipéptido de fusión modificado en la primera etapa). La primera etapa puede llevarse a cabo en un sitio, tal como en el local de la compañía proporcionando polipéptidos de fusión cuantificados, mientras que la segunda etapa se lleva a cabo en otro sitio, tal como en un laboratorio en el que se cuantifican proteínas en muestras biológicas para fines de diagnóstico.

Como se ha indicado, dichos uno o más polipéptidos diana tienen masa alterada, preferentemente están marcados con isótopo de forma diferente en comparación con dichos polipéptidos de fusión. En otras palabras, y en los casos en los que dichos polipéptidos de fusión no estén marcados con isótopos, es necesario preparar una muestra en la que el o los polipéptidos diana comprendidos en la muestra están marcados con isótopos. Por otro lado, no surge un requisito para preparar una muestra marcada con isótopos para las realizaciones que quedan dentro del tercer aspecto en el que dichos polipéptidos de fusión están marcados con isótopo.

Puede usarse más de un polipéptido de fusión que comprende diferentes subsecuencias de un polipéptido diana en dicha muestra. Se usa después más de un polipéptido de fusión en la cuantificación de un polipéptido diana dado. Este se describe adicionalmente en los ejemplos incluidos con la presente y proporciona precisión mejorada y significación estadística.

Uno o dos marcadores pueden estar presentes en dichos polipéptidos de fusión, seleccionándose dicho marcador o dichos marcadores de un marcador de purificación y un marcador de solubilidad. Esto acepta la presencia conjunta de dos marcadores diferentes. Se describen adicionalmente posteriormente implementaciones preferidas de uno de los marcadores marcadores. Se entiende que el marcador de solubilidad se usa preferentemente como un marcador de cuantificación ("secuencia marcadora") de acuerdo con los métodos.

Dicha muestra puede comprender células y/o fluidos corporales. Dichas células pueden ser de diversos tipos o de un solo tipo. Además, las células pueden incluirse en uno o más tejidos. En la medida en que se prevén células humanas, se prefiere que dicha célula humana no se obtenga de un embrión humano, en particular no mediante métodos que impliquen la destrucción de un embrión humano. Por otro lado, están al alcance del experto en la materia células madre embrionarias humanas. En consecuencia, los métodos pueden realizarse con células madre embrionarias humanas sin ninguna necesidad de usar o destruir un embrión humano. La muestra puede comprender uno o más fluidos corporales, seleccionándose preferentemente dichos fluidos corporales de sangre, suero sanguíneo, plasma sanguíneo, leche materna, líquido cefalorraquídeo, moco, líquido peritoneal, líquido pleural, saliva, semen, sudor, lágrimas, secreción vaginal y orina.

Dicha adición puede efectuarse antes de la digestión proteolítica de los polipéptidos. Esta realización se refiere a los casos en los que la muestra para analizar comprende o consiste en células. Dicha adición se refiere a la adición de un polipéptido de fusión y un polipéptido marcador de acuerdo con el método para determinar la cantidad absoluta de un polipéptido diana, o para añadir el patrón cuantitativo de acuerdo con el método para determinar la cantidad

absoluta del o los polipéptidos diana. En cualquier caso, la adición temprana posibilita que los métodos compensen cualquier sesgo posiblemente introducido por la preparación y el procesamiento de muestras, en particular por la etapa de digestión proteolítica. Esta es una ventaja adicional en comparación con los métodos de la técnica anterior que requieren una adición tardía del patrón durante el flujo de trabajo.

5 Pueden usarse entre dos y 500 polipéptidos de fusión. Como se ha indicado anteriormente, el segundo y tercer aspecto de la invención proporcionan multiplexación. Los números preferidos de polipéptidos de fusión para usar en cada caso del método son entre dos y 200, tal como entre 2 y 100, incluyendo cualquier valor entero incluido por estos límites inferior y superior tal como 50 polipéptidos de fusión. Los ejemplos incluidos con la presente  
10 proporcionan una explicación de rendimiento excelente cuando se usan 43 polipéptidos de fusión.

Puede estar presente un marcador de solubilidad en cada uno de dichos polipéptidos de fusión. Un marcador de solubilidad preferido consiste en la secuencia de SEQ ID NO: 1. La secuencia de SEQ ID NO: 1 es particularmente ventajosa porque las secuencias obtenidas por digestión triptica del proteoma humano por un lado y de la SEQ ID  
15 NO: 1 por otro lado son distintas. En otras palabras, una digestión triptica de la secuencia SEQ ID NO: 1 produce péptidos ninguno de los cuales se obtiene de una digestión triptica del proteoma humano. Lo mismo se aplica a al menos la mayoría de péptidos obtenidos de la secuencia SEQ ID NO: 1 cuando las otras enzimas preferidas como se desvela en el presente documento se usan para digestión proteolítica.

20 En una implementación dicha subsecuencia de un polipéptido (a) consiste en de 15 a 205 aminoácidos; (b) comprende un péptido proteotípico; y/o (c) se selecciona para tener identidad de secuencia mínima con otras proteínas, excluye péptidos señal y/o excluye secuencias de regiones transmembrana. La subsecuencia es la subsecuencia de un polipéptido diana comprendida en el polipéptido de fusión de acuerdo con la presente invención. La característica (a) proporciona un intervalo de longitud preferido de dicha subsecuencia. Se desvelan en el  
25 presente documento longitudes e intervalos de longitud preferidos adicionales, en particular en la descripción del cuarto aspecto de la invención. Dicha divulgación se aplica cambiando lo que deba cambiarse a la presente realización preferida. Debe observarse que dicho intervalo de longitud está por encima del intervalo de longitud observado para péptidos tripticos. Como consecuencia, esta implementación se distingue de los métodos de la técnica anterior que hacen uso de, por ejemplo, péptidos tripticos u otros péptidos que no son susceptibles de  
30 escisión por la enzima proteolítica para usar para digestión proteolítica. Provechosamente, y como se ha indicado anteriormente, las subsecuencias en este intervalo de longitud dan lugar a una pluralidad de péptidos tras digestión proteolítica, potenciando de este modo la precisión de la cuantificación.

El término "proteotípico" como se usa en este contexto específico se refiere a péptidos que se observan  
35 frecuentemente o siempre en el espectro de masas de un polipéptido dado que comprende dicho péptido proteotípico.

De acuerdo con la parte (c), se proporcionan características adicionales que se refieren a la individualidad de dicha subfrecuencia (identidad de secuencia mínima con otras proteínas, en particular con otras proteínas de mismo  
40 proteoma) o a la manipulación y/o detección fácil (exclusión de péptidos señal y segmentos transmembrana).

Dicha cantidad absoluta conocida de dicho polipéptido marcador puede determinarse por análisis de aminoácidos. Se han descrito anteriormente en el presente documento medios preferidos y métodos de análisis de aminoácidos.

45 La presente invención proporciona un polipéptido de fusión para la cuantificación de un polipéptido diana por espectroscopia de masas, en el que: dicho polipéptido de fusión consiste en de 35 a 455 restos de aminoácidos y comprende (i) una región diana, que es un fragmento del polipéptido diana y (ii) una región marcadora, que no es un fragmento del polipéptido diana, dicha región diana consiste en de 15 a 205 restos de aminoácidos y comprende al menos dos regiones de identificación; dicha región diana consiste en de 20 a 250 restos de aminoácidos y  
50 comprende al menos dos regiones de identificación; y cada región de identificación tiene la estructura Y-Z-X<sub>4-28</sub>-Y-Z, en la que todas las Y se seleccionan de uno de (i)-(iv), en los que (i) es R o K, (ii) es Y, F, W o L, (iii) es E y (iv) es D, y cada X y cada Z son de forma independiente cualquier resto de aminoácido, siempre que los Z no sean P si los Y se seleccionan de (i)-(iii); cada región de identificación comprende al menos un resto de aminoácido que comprende un isótopo pesado; dicha región marcadora consiste en proteína de unión a albúmina (ABP) o un fragmento de la  
55 misma; y el mismo resto Y puede constituir el extremo carboxílico de una primera región de identificación y el extremo amino de una segunda región de identificación.

La invención se refiere a polipéptidos de fusión que también pueden emplearse en los métodos de acuerdo con la invención. Como a lo largo de la memoria descriptiva, el polipéptido diana puede ser cualquier polipéptido, en particular un polipéptido de origen natural en el proteoma de cualquier organismo o célula en cualquier estado. Las dos regiones comprendidas en el polipéptido de fusión de acuerdo con el cuarto aspecto de la invención se seleccionan de modo que cada una de ellas comprenda al menos dos elementos estructurales específicos denominados "regiones de identificación". Resulta importante que los aminoácidos N y C terminales de cada región de identificación se seleccionan de modo que se reconozcan por una proteasa adecuada para el protocolo de  
60 espectrometría de masas descrito en el presente documento. Los aminoácidos de (i)-(iv) se basan por lo tanto en la selectividad de las siguientes proteasas: tripsina, que escinde en el lado carboxilo de restos de arginina (R) y lisina  
65

(K) a no ser que esté seguido de prolina (P); quimotripsina, que escinde el lado carboxilo de restos de tirosina (Y), fenilalanina (F), triptófano (W) y leucina (L) a no ser que esté seguido de prolina (P); Lys-C, que escinde en el lado carboxilo de restos de lisina (K) a no ser que esté seguido de prolina (P); Glu-C, que escinde en el lado carboxilo de restos de glutamato (E) a no ser que esté seguido de prolina (P); Arg-C, que escinde en el lado carboxilo de restos de arginina (R) a no ser que esté seguido de prolina (P); y Asp-N, que escinde en el lado amino de restos de aspartato (D). Este principio de diseño de los polipéptidos de fusión asegura que, tras la digestión proteolítica, se obtienen al menos dos productos proteolíticos con masa alterada de la región diana y marcadora, respectivamente. Debe entenderse que el mismo resto Y puede constituir el extremo carboxílico de una primera región de identificación y el extremo amino de una segunda región de identificación.

La expresión general "de masa alterada" se usa en el presente documento como se ha definido anteriormente. Preferentemente, se refiere a una frecuencia de al menos un isótopo que se desvía de la frecuencia o las frecuencias de origen natural del mismo, preferentemente para la aparición exclusiva de al menos un isótopo pesado, seleccionándose preferentemente isótopos pesados de D, <sup>13</sup>C y <sup>15</sup>N.

Se describe que dicha región marcadora o dicho polipéptido marcador, respectivamente, corresponde a, es decir comprende o consiste en un marcador de solubilidad o un fragmento del mismo, seleccionándose dicho marcador de solubilidad de proteína de unión a maltosa (MBP), glutatión-S-transferasa (GST), tiorredoxina (Trx), sustancia de utilización de N (NusA), modificador de ubiquitina pequeño (SUMO), un marcador potenciador de solubilidad (SET), una proteína formadora de disulfuro C (DsbC), proteína de diecisiete kilodalton (Skp), proteína quinasa T7 de fago (T7PK), dominio B1 de proteína G (GB1), dominio de repetición ZZ de IgG de proteína A (ZZ) y proteína de unión a albúmina (ABP). Las estructuras de estos marcadores de solubilidad se conocen en la técnica y están fácilmente disponibles para los expertos en la materia. Se deduce a partir de la definición anterior que el marcador de solubilidad (o fragmento del mismo) tiene masa alterada cuando constituye la región marcadora del polipéptido de fusión del cuarto aspecto.

De acuerdo con la invención, dicha región marcadora o dicho polipéptido marcador, respectivamente, comprende o consiste en proteínas de unión a albúmina (ABP) o un fragmento de la misma.

Preferentemente, dicho fragmento se elige de modo que las propiedades que confieren solubilidad se conserven o no se comprometan significativamente. El experto en la materia puede determinar si esto sucede sin más dilación, por ejemplo, realizando ensayos de solubilidad con respecto a construcciones de fusión que comprenden un polipéptido de ensayo por un lado y el marcador de solubilidad en cuestión o un fragmento del mismo por otro lado. Comparando la solubilidad de construcciones que comprenden el marcador de solubilidad completo con construcciones que comprenden un fragmento del mismo, se puede determinar si y en qué medida las propiedades que confieren solubilidad se conservan por el fragmento que se considera.

Por razones analizadas anteriormente, las secuencias de las al menos dos regiones de identificación de la región marcadora son, de acuerdo con una realización, distintas de cualquier secuencia derivable del proteoma humano por medio de proteólisis.

El polipéptido de fusión de la invención puede usarse por ejemplo en un diagnóstico de una condición médica en un sujeto que comprende la cuantificación *ex vivo* de un polipéptido diana en una muestra del sujeto. Siempre que se analicen muestras humanas, puede ser beneficioso si la región marcadora no es un polipéptido humano. Por lo tanto, en una realización del cuarto aspecto, la secuencia de aminoácidos de la región marcadora no es una secuencia de aminoácidos de una proteína humana o un fragmento de la misma. Como las proteínas humanas pueden tener alta homología con proteínas de otros eucariotas, puede preferirse particularmente si la región marcadora tiene la secuencia de aminoácidos de una proteína procariota (por ejemplo bacteriana) o un fragmento de la misma.

Como ya se ha indicado anteriormente, una región marcadora o un polipéptido marcador particularmente preferidos tienen la secuencia expuesta en SEQ ID NO: 1.

De acuerdo con realizaciones preferidas adicionales, dicha región marcadora consiste en de 40 a 150 aminoácidos, e independientemente dicha región diana consiste en de 20 a 150 aminoácidos, tal como de 25 a 100 aminoácidos. Además, se prefiere que el polipéptido de fusión consista en 80 a 300, más preferentemente de 100 a 200 aminoácidos.

De acuerdo con realizaciones preferidas adicionales, dicha región diana, y de forma independiente dicha región marcadora, comprende al menos 3, tal como al menos 4, 5, 6, 7 u 8 regiones de identificación. Estas realizaciones preferidas proporcionan un número creciente de productos proteolíticos para formar a partir de cada una de dichas regiones cuando dicho polipéptido de fusión se pone en contacto con una enzima proteolítica, detallándose adicionalmente enzimas proteolíticas posteriormente.

De acuerdo con una realización preferida adicional, cada región de identificación comprende independientemente al menos 2, tal como al menos 3 o 4 restos de aminoácidos que comprenden un isótopo pesado.

Se ha descubierto que LysC y tripsina son enzimas proteolíticas particularmente adecuadas (véase por ejemplo los ejemplos posteriores). De acuerdo con una realización preferida adicional, dichas Y se seleccionan por lo tanto de R y K.

5 Como se ha indicado anteriormente, los isótopos pesados preferidos deben seleccionarse de deuterio (D),  $^{13}\text{C}$  y  $^{15}\text{N}$ .

Normalmente, los restos de aminoácidos que comprenden un isótopo pesado del polipéptido de fusión comprenden más de un isótopo pesado. Puede preferirse un número mayor de isótopos pesados incorporados, ya que proporciona un mayor desplazamiento de masa. En una realización preferida adicional, el al menos un resto de aminoácido que comprende un isótopo pesado se selecciona de L-arginina- $^{13}\text{C}_6$ , L-arginina- $^{13}\text{C}_6$  $^{15}\text{N}_4$ , L-arginina- $^{13}\text{C}_6$  $^{15}\text{N}_4\text{D}_7$ , L-arginina- $^{15}\text{N}_4\text{D}_7$ , L-arginina- $^{15}\text{N}_4$ , L-lisina- $^{13}\text{C}_6$  $^{15}\text{N}_2$ , L-lisina- $^{15}\text{N}_2$ , L-lisina- $^{13}\text{C}_6$ , L-lisina- $^{13}\text{C}_6$  $^{15}\text{N}_2\text{D}_9$ , L-lisina- $^{15}\text{N}_2\text{D}_9$ , L-lisina-D<sub>4</sub>, L-metionina- $^{13}\text{CD}_3$ , L-tirosina- $^{13}\text{C}_9$ , L-tirosina- $^{15}\text{N}$  y L-tirosina- $^{13}\text{C}_9$  $^{15}\text{N}$ . Dichos aminoácidos marcados con isótopos pesados se conocen bien en la técnica y están disponibles de una diversidad de fabricantes. El uso de uno o más de estos aminoácidos se prefiere para cualquier esquema de marcaje y patrones de acuerdo con la presente invención. En un modo preferido, todas las lisinas y argininas se marcan de modo que los péptidos trípticos contienen típicamente un aminoácido marcado ya que la tripsina escinde específicamente en dirección C terminal de arginina y lisina.

De acuerdo con una realización preferida adicional, el polipéptido de fusión comprende además un marcador de purificación.

Además, para permitir una expresión eficaz del polipéptido de fusión, se prefiere que la región diana del polipéptido de fusión no corresponda a una región transmembrana del polipéptido diana. Además, también se prefiere que la región diana del polipéptido de fusión no corresponda a un péptido señal del polipéptido diana, ya que los péptidos señal con frecuencia se escinden en una versión madura del polipéptido diana.

En una realización preferida de cualquiera de los métodos de acuerdo con la invención como se ha descrito anteriormente, dicho polipéptido o dichos polipéptidos de fusión son como se ha definido de acuerdo con el cuarto aspecto de la presente invención así como realizaciones que se refieren al mismo.

Los marcadores de purificación preferidos deben seleccionarse de marcador His, marcador FLAG, marcador SBP, marcador myc y marcador OneStrep.

Para un usuario que cuantifica una o más proteínas o polipéptidos diana de una muestra de acuerdo con la presente divulgación, puede ser conveniente obtener el polipéptido o los polipéptidos de fusión necesarios para la cuantificación precargada en una fase sólida adecuada para la digestión proteolítica. Dicha fase sólida puede ser un soporte sólido, una columna o un filtro. Preferentemente, la cantidad de polipéptidos de fusión en dicho soporte en la columna está predeterminada. Por lo tanto, la etapa de añadir a la muestra el polipéptido o los polipéptidos de fusión no está en manos del usuario, lo que también reduce el riesgo de error humano en el procedimiento. La presente invención se refiere por tanto además a una columna en o sobre la que se dispone al menos un polipéptido de fusión de acuerdo con el cuarto aspecto. Los medios de disposición están dentro de la experiencia del experto en la materia e incluyen unión covalente así como adsorción o absorción no covalente.

Una enzima proteolítica tal como tripsina, quimotripsina, Lys-C, Glu-C o Asp-N también puede disponerse en o sobre la columna. Cuando se usa dicha columna, el usuario no tiene que añadir la enzima proteolítica para la digestión, lo que puede ser conveniente y reducir adicionalmente el riesgo de error humano. De acuerdo con una realización, el polipéptido o los polipéptidos de fusión se separan de la enzima proteolítica en el soporte/en la columna para prevenir cualquier digestión proteolítica antes de añadirse la muestra.

La presente invención proporciona además un kit que comprende: (a) al menos un polipéptido de fusión de acuerdo con la invención; y (b) (i) un segundo polipéptido que comprende o consiste en la misma secuencia de aminoácidos que la región marcadora como se define de acuerdo con el polipéptido de fusión de acuerdo con la invención pero que tiene marcaje isotópico diferente en comparación con dicha región marcadora y/o (ii) una enzima proteolítica, tal como tripsina, quimotripsina, Lys-C, Glu-C o Asp-N. La combinación de los productos necesarios para el protocolo de cuantificación descrito en el presente documento en un kit puede proporcionar reproducibilidad aumentada y riesgo reducido de error humano por parte de los usuarios. El segundo polipéptido del sexto aspecto puede estar, por ejemplo, "no marcado". También puede ser de "peso medio" o "peso extrapesado". Dichas realizaciones se han analizado anteriormente en relación con los aspectos del método.

En una realización preferida del kit, el al menos un polipéptido de fusión se dispone en o sobre una columna de acuerdo con el quinto aspecto de la invención. En una realización preferida adicional del kit, dicho segundo polipéptido se proporciona en una cantidad absoluta conocida.

En un aspecto adicional, la presente invención se refiere al uso de un polipéptido de fusión de acuerdo con la invención como referencia en una cuantificación de polipéptido diana. En una realización preferida del uso de acuerdo con la invención, dicha cuantificación se efectúa por espectrometría de masas.

Se han descrito anteriormente diversas realizaciones adicionales del aspecto de uso en relación con los otros aspectos.

Las figuras muestran:

5 **Figura 1: Flujo de trabajo esquemático para determinación precisa de concentraciones de PrEST.** ABP pesada o ligera se expresa de forma recombinante en una cepa de *E. coli* auxotrófica y se purifica usando el marcador OneStrep C terminal. La ABP marcada pesada, cuya concentración se mide por separado por análisis de aminoácidos, y el PrEST se mezclan entre sí y se realiza una digestión en solución. Los péptidos se miden con un procesamiento de CL EM/EM corto en un espectrómetro de masas de mesa y la concentración de PrEST se determina con precisión por la relación de SILAC de los péptidos de ABP que se originan del PrEST y la ABP.

10 **Figura 2: Precisión de la cuantificación de ABP.** (A) Representación de densidad de la distribución global de los 43 coeficientes de variación (CV) de los péptidos de ABP medidos en un espectrómetro de masas de mesa Exactive. (B) Proteínas ejemplares representativas que muestran las relaciones de los péptidos de ABP y los coeficientes de variación (CV).

15 **Figura 3: Relación de péptidos a lo largo de las secuencias de PrEST.** La mezcla maestra de PrEST se añadió a lisado de una línea celular de cáncer y se midió frente a la proteína endógena. Las relaciones de péptidos se extrajeron para cuantificar las proteínas. Se representa la variación de las relaciones de péptidos a lo largo de la secuencia. Los péptidos solapantes se deben a escisiones ausentes.

20 **Figura 4: Reproducibilidad del procedimiento de cuantificación absoluto.** Tres experimentos de cuantificación independientes para ejemplos representativos, en los que la preparación de mezcla maestra así como la cuantificación de PrEST se realizaron de forma independiente. Las barras reflejan la mediana de las relaciones de péptidos para cada proteína.

25 **Figura 5: Números de copias de proteínas determinados por célula de HeLa.** La representación de puntos muestra los números de copias de proteínas por célula medidos en tres experimentos independientes. Las barras de error corresponden a los CV. Se cuantificaron proteínas con números de copias que varían de 4.000 a 20.000.000 por célula (véase también Tabla 2).

30 **Figura 6: Cuantificación directa de una única proteína en lisado celular de HeLa.** (A) Principio de la estrategia "individual" para la cuantificación directa de una única proteína. En el mismo experimento, las relaciones de péptidos de SILAC que se mapean en el marcador de cuantificación de ABP determinan la cantidad de PrEST mientras que las relaciones de SILAC que se mapean en la región específica de proteína de la construcción de PrEST determinan el nivel de las proteínas endógenas. El experimento puede realizarse con células marcadas pesadas de SILAC, construcción de PrEST no marcado y marcador de ABP marcado pesado (lado izquierdo) o viceversa (lado derecho). (B) Determinación individual de cantidad de proteína absoluta. En el flujo de trabajo representado en el presente documento, una construcción de PrEST no marcada así como un marcador de ABP marcado pesado se añaden ambos a lisado celular de HeLa antes de la digestión. (C) Comparación de números de copias obtenidos del experimento de "mezcla maestra" con los de los experimentos individuales para tres proteínas diferentes. Las barras de error son desviaciones típicas de la media de mediciones por triplicado.

35 **Figura 7: Cuantificación absoluta usando PrEST pesados.** (A) Comparación de los números de copias obtenidos cuantificando PrEST ligeros frente a lisado celular pesado marcado con SILAC (símbolos negros) frente a cuantificación de PrEST pesados frente a lisado celular no marcado (símbolos rojos). (B) Valores mostrados en A pero representados como una gráfica de dispersión.

40 **Figura 8: Comparación de cuantificación basada en SILAC-PrEST y ELISA.** Se cuantificó el protooncogén c-Fos (A) y estratifina (B) mediante ELISA para evaluar la cuantificación absoluta de SILAC-PrEST. Se compararon diferentes tampones compatibles con ELISA y lisados celulares filtrados frente a no filtrados.

45 **Figura 9: Cuantificación absoluta de la integrina beta 3, talina 1 y kindlina 3 en ratones diferentes.** (a) La integrina y sus coactivadores agrupados entre sí. (b) Los niveles de expresión decrecientes de kindlina 3 en comparación con los ratones de tipo silvestre.

50 Los ejemplos ilustran la invención:

### Ejemplo 1: materiales y métodos

55 **Marcadores de identificación de epítomos proteicos** - Los fragmentos proteicos cortos, es decir las subsecuencias de polipéptidos diana, se produjeron con alto rendimiento por el Human Protein Atlas en el que se usan como antígenos para la producción de anticuerpos (18, 19). Brevemente, se diseñaron marcadores de identificación de epítomos proteicos (PrEST) adecuados que representan regiones únicas de cada proteína diana usando la secuencia del genoma humano como molde (Ensembl). Se seleccionaron PrEST únicos con un tamaño entre 50 y 150 aminoácidos y baja homología con otras proteínas humanas, incluyendo similitudes de tamaño de epítomo y dominio con otras proteínas, péptidos señal y regiones transmembrana (18). La clonación, expresión y purificación de proteínas se realizaron como se ha descrito previamente (19, 20). Para almacenamiento óptimo se liofilizaron PrEST y se disolvieron en urea 8 M y se almacenaron a -20 °C hasta su uso posterior. Para determinar que los PrEST tenían un homólogo endógeno en células HeLa, se seleccionaron 50 proteínas distribuidas sobre el intervalo de abundancia de un proteoma de HeLa que se habían medido a una profundidad de aproximadamente 4.000 proteínas. Las proteínas se seleccionaron independientemente de clases de proteínas específicas, localizaciones

celulares o funciones. De estas 50 proteínas, 43 estaban fácilmente disponibles del proceso de Protein Atlas en forma expresada de forma recombinante. Para experimentos de multiplexación estos 43 PrEST se mezclaron entre sí, cada uno a la concentración apropiada. Esta “mezcla maestra” después se añadió a lisados celulares.

5 *Cultivo celular* - Para marcaje de SILAC, se cultivaron células HeLa en medio de Eagle modificado por Dulbecco (Invitrogen) que contenía suero bovino fetal dializado al 10 % (Gibco) y penicilina/estreptomicina (Gibco). Se añadieron arginina pesada (Arg10 de alta pureza, Cambridge Isotope Laboratories) y lisina pesada (Lys8 de alta pureza, Cambridge Isotope Laboratory) a una concentración final de 33 µg/ml o 76 µg/ml, respectivamente. Después de seis pases las células se marcaron completamente como se evaluó por espectrometría de masas. Las células se  
10 contaron usando un contador celular Countess (Invitrogen) y se congelaron instantáneamente alícuotas de 106 células y se almacenaron a -80 °C.

*Expresión de proteínas de purificación de ABP (proteína de unión a albúmina)* - el vector de expresión pAff8c (Human Protein Atlas) se modificó mediante clonación de SLIC (21) insertando un marcador de afinidad OneStrep en el extremo C terminal de la proteína de unión a albúmina (ABP). Para expresar ABP marcado pesado en *E. coli*, se usó una cepa de expresión auxotrófica para arginina y lisina (33). Los cultivos se dejaron crecer en medio de autoinducción mínimo PA5052 como se ha descrito previamente en (22) pero con la adición de 18 aminoácidos normales (“ligeros”) y arginina y lisina pesada. Los cultivos se dejaron crecer durante una noche y se recogieron a una DO600 de aproximadamente 5,7. Se lisaron células *E. coli* en Tris 100 mM, NaCl 150 mM e inhibidor de proteasa (Roche) usando un Bioruptor (Diagenode). Se retiraron los residuos celulares por centrifugación y se purificó ABP soluble usando cromatografía de afinidad en una columna StrepTap Hitrap (GE Healthcare) acoplada a un sistema de ÄKTA. La pureza de la proteína se evaluó por espectrometría de masas mediante una digestión en solución seguida de CL EM/EM. Las abundancias de ABP y contaminantes se estimaron añadiendo la señal para sus péptidos más intensos. ABP se dializó en PBS, se separó en alícuotas, se congeló instantáneamente y se  
15 almacenó a -80 °C. La concentración de ABP purificada se midió por análisis de aminoácidos (Genaxxon BioScience GmbH).

*Preparación de muestras* - Se lisaron células HeLa en Tris 100 mM, SDS 4 %, DTT 100 mM, se incubaron durante 5 min a 95 °C y se rompieron usando un Bioruptor. El lisado se eliminó por centrifugación mediante filtros SpinX  
20 (22 µm, Corning). Los PrEST se añadieron a concentraciones apropiadas (véase texto principal) a células HeLa marcadas y las muestras se procesaron adicionalmente por el método de FASP (23). Brevemente, se capturaron proteínas en un filtro de 30 kDa y se intercambió SDS con un tampón que contiene urea. Las proteínas se alquilaron con yodoacetamida y se tripsinizaron (Promega). Se realizó separación de péptidos adicional usando fracción seis basada en pipeta SAX como se describe (24).

Los PrEST y ABP se mezclaron y se solubilizaron en tampón de desnaturalización (urea 6 M, tiourea 2 M en HEPES 10 mM, pH 8), se redujeron con DTT y se alquilaron posteriormente con yodoacetamida. La mezcla de proteínas se digirió con LysC (Wako) durante 3 h, se diluyó con bicarbonato de amonio y se digirió adicionalmente con tripsina durante una noche. La digestión se detuvo acidificando con TFA y se desaló en StageTips de disco Empore -C<sub>18</sub>  
35 (25).

*Cromatografía líquida y espectrometría de masas* - Se realizó análisis de los PrEST ligeros añadidos a células HeLa en un espectrómetro de masas LTQ-Orbitrap (Thermo Fisher Scientific) acoplado a una nano-HPLC Easy mediante una fuente de iones de nanoelectropulverización (Proxeon Biosystems, ahora Thermo Fisher Scientific). Los péptidos se separaron en un emisor de sílice fusionado de 15 cm empaquetado de forma interna con material de fase inversa resina ReproSil-Pur 120 C18-AQ de 3 µm (Dr. Maisch GmbH) y se eluyó con un gradiente de 205 min de tampón B 5-35 % (acetonitrilo 80 %, ácido acético 0,5 %). El espectrómetro de masas se operó de una manera dependiente de los datos para medir automáticamente EM y EM/EM consecutivas. Se adquirieron espectros de EM de exploración completa de LTQ-Orbitrap (desde 300 a 1650 m/z) con una resolución de 60.000 a m/z 400. Los siete iones más abundantes se aislaron secuencialmente y se fragmentaron en la trampa de iones lineal usando disociación inducida por colisión (CID) seguida de análisis en la trampa de iones lineal.

El análisis de los PrEST añadidos a las células HeLa se realizó en un espectrómetro de masas LTQ-Orbitrap Velos (Thermo Fisher Scientific) acoplado a una nano-HPLC Easy mediante una fuente de iones de nanoelectropulverización (Thermo Fisher Scientific). Los péptidos se separaron en una columna de 20 cm empaquetada de forma interna usando resina C18-AQ de 1,8 µm (Dr. Maisch GmbH) y se eluyeron con un gradiente de 205 min de tampón B 5-35 %. El espectrómetro de masas se operó de una manera dependiente de datos para medir automáticamente EM y 10 EM/EM consecutivas aplicando disociación de colisión de mayor energía (HCD) (34). Se adquiriendo espectros de EM de exploración completa LTQ-Orbitrap (de 100 o 300 a 1650 m/z) con una resolución de 60.000 a m/z 400.

Los péptidos PrEST-ABP se analizaron en línea con el instrumento Exactive con opción de HCD (Thermo Fisher Scientific) usando la misma preparación de nano-HPLC que se ha descrito anteriormente. Los péptidos se eluyeron con un gradiente lineal con tampón B 5-30 % durante 40 min. El espectrómetro de masas Exactive identificó péptidos con fragmentación de todos los iones (AIF) realizando exploraciones de EM alternantes (300-1.600 m/z) de los iones precursores y todas las exploraciones de fragmentación de iones (100-1.600 m/z) usando fragmentación de HCD por  
65

etapas (26). Ambas exploraciones se adquirieron a una resolución de 100.000 a m/z 200.

*Análisis de datos* - Se analizaron datos adquiridos con MaxQuant (27) (versión 1.1.1.36) usando la base de datos IPI humana (v 3.68 – 87.083 entradas). Se añadieron contaminantes habituales y la secuencia del marcador de solubilidad de ABP a esta base de datos. Para identificación de péptidos se usó Andromeda, un motor de búsqueda probabilístico incorporado en el marco de MaxQuant (28). Se incluyó la carbamidometilación de cisteína en la búsqueda como una modificación fija y se incluyeron oxidación de metionina así como acetilación N terminal como modificaciones variables. Se permitieron dos escisiones erróneas y se requirió un mínimo de seis aminoácidos por péptido identificado. La tolerancia de masa inicial para iones precursores o iones de fragmentos se estableció en 6 ppm y se permitió que las masas de fragmentos se desviaran hasta 0,5 Th. Para evaluación estadística de los datos obtenidos, se usaron la probabilidad de error posterior y tasa de falso descubrimiento (FDR). La FDR se determinó buscando en una base de datos inversa y se estableció a 0,01 para identificación de péptidos.

Los datos de AIF se procesaron como se ha descrito anteriormente excepto que se analizaron hasta 50 picos por 100 m/z con una tolerancia de 15 ppm. La masa de ion precursor se igualó a los candidatos de iones de fragmentos posibles basándose en el valor de correlación de coseno de al menos 0,6 (26).

*Ensayo inmunoabsorbente ligado a enzimas* - Se llevaron a cabo mediciones de cantidades absolutas de protooncogén c-Fos y estratifina (14-3-3  $\sigma$ ) por ELISA. Los kits se obtuvieron de USCNK Life Science y se realizaron de acuerdo con las instrucciones del fabricante. Las células HeLa se lisaron en PBS, RIPA 1 (Tris 50 mM pH 7,5, NaCl 150 mM, NP-40 1 %) o RIPA 2 (Tris 50 mM pH 7,5, NaCl 150 mM, NP-40 1 %; SDS 0,1 %) con inhibidores de proteasa. Las células se rompieron por 3 ciclos de congelación-descongelación y sonicación usando el Bioruptor. Para el ELISA las muestras se diluyeron 1:10. La actividad de fluorescencia se midió por un lector de microplacas (Tecan) y se convirtió a concentración real por una curva patrón.

## Ejemplo 2: cuantificación absoluta de proteínas en células HeLa

A diferencia de la cuantificación relativa, puede efectuarse cuantificación absoluta como un proceso de dos etapas que requiere en primer lugar la medición de la cantidad absoluta del patrón y en segundo lugar la cantidad relativa del patrón en comparación con el analito de interés. La determinación y el control posterior del nivel del patrón no es de ninguna manera trivial y puede ser fácilmente la etapa que limite la precisión global del enfoque. Posteriormente, se describe en primer lugar un método genérico para determinar la cantidad absoluta de cada PrEST con alta precisión.

Después se construye una “mezcla maestra” de diferentes PrEST y se evalúa la capacidad del método de SILAC - PrEST para cuantificar con precisión proteínas celulares. Después se aplica la mezcla maestra para determinar los números de copias de 37 proteínas en una línea celular de cáncer. Finalmente, se describe un flujo de trabajo alternativo para la cuantificación de proteínas individuales de interés, en el que las dos etapas se combinan en un análisis de CL EM/EM.

*Medición precisa de concentraciones de PrEST* - Cada PrEST ya está fusionado con el dominio de unión a albúmina (ABP), un marcador de solubilización de 120 aminoácidos. La digestión por ordenador de ABP da como resultado 40 péptidos dípticos con una longitud de entre 6 y 30 aminoácidos (Tabla Complementaria 1). Se expresó de forma recombinante una versión marcada con SILAC pesado del marcador proteico ABP. Cuando fue necesario, se usó un enfoque de afinidad doble basado en un marcador His N terminal y un marcador OneStrep C terminal para generar fragmento proteico altamente purificado y para asegurar que solamente se obtenía ABP de longitud completa. La concentración absoluta de fragmento de proteína ABP se determinó por análisis de aminoácidos, que es el método más preciso para cuantificación de proteínas, pero que solamente es aplicable a proteínas altamente purificadas en cantidades relativamente grandes. La incorporación de SILAC pesado en ABP fue del 99 % y su pureza fue de aproximadamente el 97 % como se valoró por espectrometría de masas (véase procedimientos experimentales). Debido a que estos dos factores actúan en una dirección compensatoria y debido al pequeño tamaño del efecto, la concentración medida de ABP no se ajustó para ellos.

**Tabla Complementaria 1: todos los péptidos ABP detectados en los ciclos de AIF.** Todos los péptidos por ordenador del marcador de solubilidad ABP así como los péptido identificados cuando se determinó la concentración precisa del PrEST (véase Figura 1) para las tres mezclas maestras.

Secuencia peptídica	Longitud	Masa	Escisión ausente	Detectado
TVEGVK	6	631,354	0	X
NLINNAK	7	785,440	0	X
SIELAEAK	8	859,465	0	X
YGVSDYHK	8	967,440	0	X
YGVSDYYK	8	993,444	0	X
VLANRELDK	9	1056,593	1	X
SQTPAEDTVK	10	1074,519	0	X

Secuencia peptídica	Longitud	Masa	Escisión ausente	Detectado
DLQAQWESAK	11	1186,619	0	X
GSHMASIAEAK	11	1100,528	0	
DLQAQWESAKK	12	1314,714	1	X
EIDKYGVSDYHK	12	1452,689	1	X
ELDKYGVSDYYK	12	1478,693	1	X
ISEATDGLSDFLK	13	1394,693	0	X
NLINNAKTVEGVK	13	1398,783	1	
SIELAEAKVLANR	13	1412,799	1	
DLQAQWESAKKAR	14	1541,853	2	
ARISEATDGLSDFLK	15	1621,831	1	X
YGVSDYHKNLINNAK	15	1734,869	1	
YGVSDYYKNLINNAK	15	1760,873	1	
GSHMASLAEAKVLANR	16	1653,862	1	
KARISEATDGLSDFLK	16	1749,926	2	X
MGSSHHHHHSSGLVPR	17	1898,882	0	
SIELAEAKVLANRELDK	17	1898,047	2	X
TVEGVKDLQAQWESAK	17	1799,963	1	X
VLANRELDKYGVSDYHK	17	2006,022	2	X
VLANRELDKYGVSDYYK	17	2032,027	2	X
SQTPAEDTVKSIELAEAK	18	1915,974	1	X
TVEGVKDLQAQVVESAKK	18	1928,058	2	X
ELDKYGVSDYHKNLINNAK	19	2220,118	2	
ELDKYGVSDYYKNLINNAK	19	2246,122	2	
GGGSGGGSSGSAWSHPQFEK	20	1845,803	0	
GSHMASLAEAKVLANRELDK	20	2139,111	2	
YGVSDYHKNLINNAKTVEGVK	21	2348,213	2	
YGVSDYYKNLINNAKTVEGVK	21	2374,217	2	
ISEATDGLSDFLKSQTPAEDTVK	23	2451,202	1	X
SQTPAEDTVKSIELAEAKVLANR	23	2469,308	2	
NLINNAKTVEGVKDLQAQWESAK	24	2567,392	2	
ARISEATDGLSDFLKSQTPAEDTVK	25	2678,34	2	X
AUDEILAALPGTFAHYGSAWSHPQFEK	28	3068,54	0	
MGSSHHHHHSSGLVPRGSHMASLAEAK	28	2981,4	1	

- CL EM/EM de ABP reveló de hecho muchos péptidos tripticos fácilmente detectables (véase posteriormente). Cada uno de los 43 PrEST del proyecto de Protein Atlas se mezcló por separado con una cantidad conocida de ABP marcado como se perfila esquemáticamente en la Figura 1 para permitir un experimento de CL- EM/EM de SILAC.
- 5 Como este experimento requiere un ciclo de CL EM/EM separado para cada PrEST era probable que fuera limitante de la velocidad para el proyecto general.
- Se decidió por lo tanto realizar este análisis en un instrumento Orbitrap de sobremesa robusto y económico en lugar de un instrumento Velos. El instrumento Exactive no puede aislar precursores peptídicos, por lo tanto se identificaron los péptidos por la fragmentación de todos los iones (AIF) (26) en ciclos de 1 h. Típicamente se pudieron cuantificar al menos ocho péptidos ABP marcados frente a los péptidos ABP correspondientes de los PrEST, lo que condujo a un coeficiente de variación (CV) mediano de 7 % para cuantificación de PrEST (Figura 2A).
- 10
- Para superar la etapa de medición de la concentración de PrEST, que limita el rendimiento general, los PrEST pesados se midieron por nanoelectropulverización estática en un sistema basado en microplaca automático (TriVersa Nanomate). Esto permitió mayores mediciones de rendimiento de estas mezclas sencillas de péptidos de ABP usando consumo de muestras bajo. La relación de péptidos mostró un coeficiente mediano de variación 5,5 %, una mejora sobre la medida basada en Exactive de 7 %.
- 15
- Es importante que pueda repetirse una cuantificación de PrEST particular en este estadio hasta que se consiga una precisión deseada. Aquí, esto no se realizó, ya que se estimó que la precisión de cuantificación de PrEST era mayor que la de las otras etapas en el flujo de trabajo. Se muestran algunos ejemplos típicos de resultados de la cuantificación de PrEST en la Figura 2B. Obsérvese que la precisión de cuantificación no depende de la abundancia celular o ningún otro atributo de la proteína diana, ya que se usan las mismas cantidades de PrEST en cada
- 20

experimento de cuantificación de PrEST. Resulta importante que la precisión de cuantificación en el flujo de trabajo tampoco dependa de la pureza del PrEST porque el método de los inventores mide específicamente la concentración de PrEST y no de proteína total.

5 *Mezcla maestra de PrEST y cuantificación de proteína endógena* - Habiendo cuantificado las cantidades de PrEST se procedió a medir los niveles de expresión de proteínas en una línea celular de cáncer humano. Por conveniencia se usaron PrEST no marcados y se cuantificaron frente a células HeLa marcadas con SILAC pesado. Ya que los lisados celulares totales digeridos consisten en cientos de miles de péptidos trípticos, la adición de un único PrEST o incluso un gran número de PrEST no cambia la complejidad general de la mezcla. Basándose en las cantidades  
10 cuantitativas establecidas anteriormente, se han mezclado aquí 43 PrEST entre sí. En experimentos iniciales se usaron mezclas equimolares de PrEST, que se añadieron a lisado de HeLa en diferentes cantidades. Las relaciones de SILAC medidas establecieron niveles apropiados de cada PrEST en la mezcla maestra, de modo que las relaciones de SILAC estaban dentro del intervalo cuantificable de forma más precisa, es decir relativamente cerca entre sí.

15 La mezcla maestra con niveles apropiados de los 43 PrEST se añadió al lisado de células marcadas con SILAC. La mezcla se digirió de acuerdo con el protocolo de FASP seguido de fraccionamiento de SAX y dando como resultado seis fracciones que se midieron por separado con gradientes de 4 h en un espectrómetro de masas LTQ Orbitrap. Se pudieron cuantificar 37 de las 43 proteínas a las que se dirigía la mezcla maestra de PrEST de los inventores.

20 Las proteínas se cuantificaron en general con varios péptidos derivados de PrEST (promedio 3,7 y mediana 3), lo que condujo a un CV mediano general de 18 % (Tabla Complementaria 2). Los resultados para estas 37 dianas proteicas se muestran en la Figura 3 y se describe la identificación completa e información de cuantificación en la Tabla Complementaria 2. Como ejemplo, la proteína de adhesión IQGAP1 se cuantificó con cinco péptidos, que  
25 proporcionaron cada uno resultados de cuantificación casi idénticos (CV 10,6 %). Seis de los siete péptidos trípticos cuantificados de ATP5B (subunidad beta de ATP sintasa mitocondrial), tuvieron relaciones de SILAC muy cercanas, sin embargo, un péptido tenía una relación que difería el 38 % de la mediana. Este péptido es claramente un valor atípico y su valor de desviación contribuye sustancialmente al valor de CV, elevándolo del 8,2 % al 27,2 %. Obsérvese, sin embargo, que se basa la cuantificación de proteínas en la mediana de los valores peptídicos; por lo tanto el péptido atípico apenas contribuye al valor de expresión de proteína medido y el valor de CV por lo tanto  
30 subestima la precisión de hecho obtenida en este experimento. Por la misma razón las modificaciones de las proteínas endógenas en la región abarcada por el PrEST podrían provocar relaciones de péptidos atípicos, lo que contribuiría poco a la relación de proteína medida.

**Tabla Complementaria 2:** toda la información de identificación y cuantificación usada para cuantificar proteínas.

Nombres de proteínas	Nombre delSecuencia gen	Relación de H/L de mezcla maestra		Relación de H/L de mezcla maestra		Relación de H/L de mezcla maestra	
		(1)	CV (%)	(2)	CV (%)	(3)	CV (%)
Acil coenzima citosólica	ACOT7	NaN	-	NaN	-	NaN	-
Acil coenzima citosólica	ACOT7	NaN	-	NaN	-	NaN	-
Acil coenzima citosólica	ACOT7	0,27622		NaN	0,25644	NaN	5,2343
Acil coenzima citosólica	ACOT7	NaN		NaN		NaN	
Proteína de tipo AFG3 2	AFG3L2	0,77684	6,27	NaN	0,82239	NaN	1,4433
Proteína de tipo AFG3 2	AFG3L2	0,68028		NaN	0,78043	NaN	1,3793
Proteína de tipo AFG3 2	AFG3L2	0,80176		NaN	0,82107	NaN	
Proteína de tipo AFG3 2	AFG3L2	0,72983		NaN	0,87038	NaN	1,5475
Proteína de tipo AFG3 2	AFG3L2	0,76345		NaN	0,87166	NaN	
Proteína de tipo AFG3 2	AFG3L2	0,12868		NaN	0,12046	NaN	0,87365
Dominio AAA de familia de ATPasa	ATAD2	0,82979	23,09	NaN	0,64229	NaN	1,0596
Subunidad beta de ATP sintasa	ATP5B	0,85299		NaN	0,84263	NaN	1,1617
Subunidad beta de ATP sintasa	ATP5B	0,71767		NaN	0,73297	NaN	0,90287
Subunidad beta de ATP sintasa	ATP5B	0,4194		NaN	0,76715	NaN	1,1006
Subunidad beta de ATP sintasa	ATP5B	0,76219		NaN	0,67515	NaN	0,99543
Subunidad beta de ATP sintasa	ATP5B	0,89528		NaN	0,95159	NaN	1,1652
Proteína de dedo de zinc 828	C13orf8	NaN	65,67	NaN	43,02	NaN	65,84
Proteína de dedo de zinc 828	C13orf8	NaN		NaN		NaN	
Proteína de dedo de zinc 828	C13orf8	NaN		NaN		NaN	
Proteína de dedo de zinc 828	C13orf8	0,081796		NaN	0,099583	NaN	0,34653
Proteína de dedo de zinc 828	C13orf8	NaN		NaN		NaN	
Proteína de dedo de zinc 828	C13orf8	NaN		NaN		NaN	
Proteína de dedo de zinc 828	C13orf8	0,22363		NaN	0,18665	NaN	0,9503
Interacción de tallo- bucle de SRA	C14orf156	3,5457	74,19	NaN	2,9157	NaN	1,466
Interacción de tallo- bucle de SRA	C14orf156	0,35422		NaN	2,9391	NaN	1,5112
Interacción de tallo- bucle de SRA	C14orf156	NaN		NaN	0,33818	NaN	0,2253
Interacción de tallo- bucle de SRA	C14orf156	NaN		NaN		NaN	1,6141
Interacción de tallo- bucle de SRA	C14orf156	3,5138		NaN	3,2266	NaN	1,5357
Interacción de tallo- bucle de SRA	C14orf156	NaN		NaN	0,29812	NaN	1,6403

Tabla complementaria 2 continuación

µl PrESTs (1+2)	µl PrESTs (3)	Exactive pmol/µl 1	Exactive pmol/µl 2	Exactive pmol/µl 3	PrEST 1 pmol	PrEST 2 pmol	PrEST 3 pmol	Proteína 1 pmol	Proteína 2 pmol	Proteína 3 pmol
0,011	0,0017	275,36565	273,155	94,187082	3,029022	3,004705	0,160118	#VALUE!	#VALUE!	#VALUE!
0,011	0,0017	275,36565	273,155	94,187082	3,029022	3,004705	0,160118	#VALUE!	#VALUE!	#VALUE!
0,011	0,0017	275,36565	273,155	94,187082	3,029022	3,004705	0,160118	0,836676	0,770527	0,838106
0,011	0,0017	275,36565	273,155	94,187082	3,029022	3,004705	0,160118	#VALUE!	#VALUE!	#VALUE!
0,011	0,0017	275,36565	273,155	94,187082	3,029022	3,004705	0,160118	#VALUE!	#VALUE!	#VALUE!
0,006	0,0107	125,05216	129,519	17,806094	0,750313	0,777114	0,190525	0,582873	0,639091	0,274985
0,006	0,0107	125,05216	129,519	17,806094	0,750313	0,777114	0,190525	0,510423	0,606483	0,262791
0,006	0,0107	125,05216	129,519	17,806094	0,750313	0,777114	0,190525	0,601571	0,638065	#VALUE!
0,006	0,0107	125,05216	129,519	17,806094	0,750313	0,777114	0,190525	0,547601	0,676384	0,294838
0,006	0,0107	125,05216	129,519	17,806094	0,750313	0,777114	0,190525	0,572826	0,677379	#VALUE!
0,005	0,0025	179,90794	184,773	70,124349	0,89954	0,923865	0,175311	0,115753	0,111289	0,15316
0,049	0,0855	239,56298	196,145	81,16038	11,73859	9,611105	6,939212	9,740561	6,173117	7,35279
0,049	0,0855	239,56298	196,145	81,16038	11,73859	9,611105	6,939212	10,0129	8,098605	8,061283
0,049	0,0855	239,56298	196,145	81,16038	11,73859	9,611105	6,939212	8,424431	7,044652	6,265207
0,049	0,0855	239,56298	196,145	81,16038	11,73859	9,611105	6,939212	4,923163	7,373159	7,637297
0,049	0,0855	239,56298	196,145	81,16038	11,73859	9,611105	6,939212	8,947033	6,488938	6,9075
0,049	0,0855	239,56298	196,145	81,16038	11,73859	9,611105	6,939212	10,50932	9,145831	8,08557
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	#VALUE!	#VALUE!	#VALUE!
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	#VALUE!	#VALUE!	#VALUE!
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	#VALUE!	#VALUE!	#VALUE!
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	0,064421	0,057554	0,064176
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	#VALUE!	#VALUE!	#VALUE!
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	#VALUE!	#VALUE!	#VALUE!
0,007	0,0048	112,51212	82,565	38,582311	0,787585	0,577955	0,185195	0,176128	0,107875	0,175991
0,005	0,0378	151,61694	90,9	40,283794	0,758085	0,4545	1,522727	2,687941	1,325186	2,232318
0,005	0,0378	151,61694	90,9	40,283794	0,758085	0,4545	1,522727	#VALUE!	1,335821	2,301146
0,005	0,0378	151,61694	90,9	40,283794	0,758085	0,4545	1,522727	0,268529	0,153703	0,34307
0,005	0,0378	151,61694	90,9	40,283794	0,758085	0,4545	1,522727	#VALUE!	#VALUE!	2,457834
0,005	0,0378	151,61694	90,9	40,283794	0,758085	0,4545	1,522727	2,663758	1,46649	2,338452
0,005	0,0378	151,61694	90,9	40,283794	0,758085	0,4545	1,522727	#VALUE!	0,135496	2,49773

Tabla Complementaria 2 continuación

Nombres de proteínas	Nombre del gen	Secuencia	Relación de H/L de mezcla maestra (1)		Relación de H/L de mezcla maestra (2)		Relación de H/L de mezcla maestra (3)	
			CV (%)	H/L de mezcla maestra	CV (%)	H/L de mezcla maestra	CV (%)	H/L de mezcla maestra
Proteína no caracterizada	C1orf65	ILVELADEK	-	NaN	-	NaN	-	NaN
Carcinoma hepatocelular	C9orf78	GDSESEEDQDSEEV	-	NaN	-	NaN	-	NaN
Carcinoma hepatocelular	C9orf78	RGDSESEEDQDSEEV	-	NaN	-	NaN	-	NaN
Carcinoma hepatocelular	C9orf78	VQEETTLVDDPFQMK	0,34795	0,27853	0,27853	0,54925	0,54925	0,54925
Carbonil reductasa [NA	CBR3	AFENCSEDLQER	0,12453	0,10329	0,10329	2,5724	2,5724	-
Carbonil reductasa [NA	CBR3	FHSETLTEGDLVDMK	0,12993	NaN	NaN	NaN	NaN	NaN
Carbonil reductasa [NA	CBR3	TNFFATR	NaN	0,091036	0,091036	NaN	NaN	NaN
Carbonil reductasa [NA	CBR3	WNISSLQCLR	NaN	0,11067	0,11067	NaN	NaN	NaN
Dominio superenrollado-cont	CCDC55	NQEKPSNCESSLGAK	-	NaN	-	NaN	-	NaN
Subunidad de proteína de complejo T 1	CCT2	HGINCFINR	0,63785	0,42168	0,42168	1,3197	1,3197	7,84
Subunidad de proteína de complejo T 1	CCT2	ILANTGMDTDK	0,47498	0,37474	0,37474	1,0965	1,0965	-
Subunidad de proteína de complejo T 1	CCT2	IJANTGMDTDKIK	0,26858	0,1831	0,1831	NaN	NaN	NaN
Subunidad de proteína de complejo T 1	CCT2	LALVTGGEIASTFDHPPELVK	0,5415	0,51676	0,51676	1,2981	1,2981	-
Subunidad de proteína de complejo T 1	CCT2	LIEEVMIGEDK	0,34771	0,30789	0,30789	1,2599	1,2599	-
Subunidad de proteína de complejo T 1	CCT2	VAEIEHAEK	0,4707	0,45578	0,45578	1,1598	1,1598	-
Subunidad de proteína de complejo T 1	CCT2	VAEIEHAEKEK	0,51512	0,51219	0,51219	1,3444	1,3444	-
Multivesicular con carga	CHMP6	IAQQLER	0,12113	0,07044	0,07044	NaN	NaN	-
Multivesicular con carga	CHMP6	YQEQLLDR	-	0,080177	0,080177	NaN	NaN	-
Comp de signalosoma COP9	COPS5	DHHYFK	0,9528	1,1093	1,1093	1,7291	1,7291	11,08
Comp de signalosoma COP9	COPS5	ISALALLK	0,90507	0,82773	0,82773	1,9194	1,9194	-
Comp de signalosoma COP9	COPS5	SGNLEVMGLMLGK	0,96904	1,2254	1,2254	2,157	2,157	-
Comp de signalosoma COP9	COPS5	VDGETMIIMDSFALPVEGTE	NaN	NaN	NaN	NaN	NaN	NaN
Citocromo b5 reductasa	CYB5R4	LLHDLNFSK	-	NaN	-	NaN	NaN	34,50
Citocromo b5 reductasa	CYB5R4	QGHISPALLSEFLK	-	NaN	-	NaN	NaN	0,077032
Citocromo b5 reductasa	CYB5R4	TEDDIWR	0,035422	0,034486	0,034486	0,12675	0,12675	-
Dependiente de ATP probable	DDX20	GEENMMMR	0,60662	0,59133	0,59133	21,50	21,50	-
Dependiente de ATP probable	DDX20	VLISTDLTSR	-	0,43526	0,43526	NaN	NaN	-

Tabla complementaria 2 continuación

µl PrESTs (1+2)	µl PrESTs (3)	Exactive pmol/µl 1	Exactive pmol/µl 2	Exactive pmol/µl 3	pmol PrEST 1	pmol PrEST 2	pmol PrEST 3	pmol Proteina 1	pmol Proteina 2	pmol Proteina 3
0,006	0,0026	221,44221	130,904	52,045542	1,328653	0,785424	0,135318	#VALUE!	#VALUE!	#VALUE!
0,006	0,0103	242,34528	179,229	78,04384	1,454072	1,075374	0,803852	#VALUE!	#VALUE!	#VALUE!
0,006	0,0103	242,34528	179,229	78,04384	1,454072	1,075374	0,803852	#VALUE!	#VALUE!	#VALUE!
0,006	0,003	164,39734	155,761	70,484927	0,986384	0,934566	0,211455	0,505944	0,299524	0,441515
0,006	0,003	164,39734	155,761	70,484927	0,986384	0,934566	0,211455	0,122834	0,096531	0,543946
0,006	0,003	164,39734	155,761	70,484927	0,986384	0,934566	0,211455	0,128161	#VALUE!	#VALUE!
0,006	0,003	164,39734	155,761	70,484927	0,986384	0,934566	0,211455	#VALUE!	0,085079	#VALUE!
0,015	0,0107	86,173017	86,118	38,073055	1,292595	1,29177	0,407382	#VALUE!	0,103428	#VALUE!
0,077	0,0951	337,27256	140,659	61,125623	25,96999	10,83074	5,813047	16,56496	4,567108	7,671478
0,077	0,0951	337,27256	140,659	61,125623	25,96999	10,83074	5,813047	12,33522	4,058713	6,374006
0,077	0,0951	337,27256	140,659	61,125623	25,96999	10,83074	5,813047	6,975019	1,983109	#VALUE!
0,077	0,0951	337,27256	140,659	61,125623	25,96999	10,83074	5,813047	14,06275	5,596895	7,545916
0,077	0,0951	337,27256	140,659	61,125623	25,96999	10,83074	5,813047	9,030024	3,334677	7,323858
0,077	0,0951	337,27256	140,659	61,125623	25,96999	10,83074	5,813047	12,22407	4,936436	6,741972
0,011	0,0053	151,27677	86,581	38,633434	1,664044	0,952391	0,204757	0,201566	0,067086	7,81506
0,011	0,0053	151,27677	86,581	38,633434	1,664044	0,952391	0,204757	#VALUE!	0,07636	#VALUE!
0,004	0,0119	129,90234	97,939	31,757433	0,519609	0,391756	0,377913	0,495084	0,434575	0,65345
0,004	0,0119	129,90234	97,939	31,757433	0,519609	0,391756	0,377913	0,470283	0,324268	0,725367
0,004	0,0119	129,90234	97,939	31,757433	0,519609	0,391756	0,377913	0,503522	0,480058	0,815159
0,004	0,0119	129,90234	97,939	31,757433	0,519609	0,391756	0,377913	#VALUE!	#VALUE!	#VALUE!
0,006	0,0044	133,16874	85,926	35,243385	0,799012	0,515556	0,155071	#VALUE!	#VALUE!	#VALUE!
0,006	0,0044	133,16874	85,926	35,243385	0,799012	0,515556	0,155071	#VALUE!	#VALUE!	#VALUE!
0,006	0,0044	133,16874	85,926	35,243385	0,799012	0,515556	0,155071	0,028303	0,017779	0,019655
0,005	0,0039	126,07369	113,423	42,562929	0,630368	0,567115	0,165995	0,382394	0,335352	#VALUE!
0,005	0,0039	126,07369	113,423	42,562929	0,630368	0,567115	0,165995	#VALUE!	0,246842	#VALUE!

Tabla complementaria 2 continuación

Nombres de proteínas	Nombre gen	Secuencia del	Relación de H/L de mezcla maestra		Relación de H/L de mezcla maestra		Relación de H/L de mezcla maestra	CV (%)	CV (%)	Relación de H/L de mezcla maestra	CV (%)
			(1)	(2)	(2)	(3)					
Enoil-CoA hidratasa, m	ECHS1	EGMTAFVEK	0,19635	13,95	0,17962	14,70	1,3826	9,67	1,3826	9,67	
Enoil-CoA hidratasa, m	ECHS1	ESVNAAFEMTLTEGSK	0,14056		0,13711		1,0823		1,0823		
Enoil-CoA hidratasa, m	ECHS1	ICPVETLVEEAIQCAEK	0,22122		0,19149		1,3474		1,3474		
Enoil-CoA hidratasa, m	ECHS1	ISAQDAK	NaN		NaN		NaN		NaN		
Enoil-CoA hidratasa, m	ECHS1	IVVAMAK	0,15532		0,16412		NaN		NaN		
Enoil-CoA hidratasa, m	ECHS1	KEGMTAFVEK	NaN		NaN		NaN		NaN		
Enoil-CoA hidratasa, m	ECHS1	KLFYSTFATDDR	NaN		0,13372		NaN		NaN		
Enoil-CoA hidratasa, m	ECHS1	LFYSTFATDDR	0,16792		0,15756		1,1803		1,1803		
Enoil-CoA hidratasa, m	ECHS1	LFYSTFATDDR	0,18416		0,20966		1,1801		1,1801		
Enoil-CoA hidratasa, m	ECHS1	QAGLVSK	0,17537		0,16861		1,2314		1,2314		
Enoil-CoA hidratasa, m	ECHS1	SLAMEMVLTGDR	0,17505		0,18633		1,0857		1,0857		
Traducción eucariota ir	EIF3E	LGHVVMGNNAVSPYQQVIEK	3,4941	5,39	1,4287	19,18	1,2643	7,57	1,2643	7,57	
Traducción eucariota ir	EIF3E	LNMTPEEAER	NaN		1,877		1,431		1,431		
Traducción eucariota ir	EIF3E	SQMLAMNIEK	3,2375		NaN		NaN		NaN		
Traducción eucariota ir	EIF3E	WVNLIR	NaN		NaN		NaN		NaN		
Reticulo endoplásmico I	ERLIN2	ADAEQCYTAMK	0,42437	8,66	0,39751	10,66	1,253	18,18	1,6231	18,18	
Reticulo endoplásmico I	ERLIN2	DIPNMFIDSAGSVSK	0,37141		0,35448		1,474		1,474		
Reticulo endoplásmico I	ERLIN2	LSFGLDEPLETATK	0,34781		0,35101		0,99961		0,99961		
Reticulo endoplásmico I	ERLIN2	LTPEYLQLMK	0,43322		0,44968		1,6324		1,6324		
Reticulo endoplásmico I	ERLIN2	QFEGLDK	0,37105		0,41697		1,5783		1,5783		
Reticulo endoplásmico I	ERLIN2	VAQVAEITYGQK	0,40674		NaN		NaN		NaN		
Ácido graso sintasa	FASN	AALQEELQLCK	0,87952	13,38	0,79607	12,46	1,0964	15,17	1,0964	15,17	
Ácido graso sintasa	FASN	DPSQQQLPR	0,80543		NaN		NaN		NaN		
Ácido graso sintasa	FASN	FCFTPHTEEGCLSER	0,79217		0,74449		1,0216		1,0216		
Ácido graso sintasa	FASN	GLVQALQTK	0,91099		0,69768		1,262		1,262		
Ácido graso sintasa	FASN	LLSAACR	1,0665		NaN		NaN		NaN		
Ácido graso sintasa	FASN	MVVPGLDGAQIPR	0,77952		0,60727		1,1329		1,1329		
Ácido graso sintasa	FASN	QQEQQVPILEK	0,73946		0,69517		0,8985		0,8985		
Ácido graso sintasa	FASN	RQQEQQVPILEK	0,69837		0,67562		0,87805		0,87805		
Ácido graso sintasa	FASN	VTTQQGLK	0,83488		0,88813		1,1698		1,1698		
Ácido graso sintasa	FASN	VTVAGGVHISGLHTESAPR	0,71007		0,64471		0,82048		0,82048		

Tabla complementaria 2 continuación

µl PrESTs (1+2)	µl PrESTs (3)	Exactive pmol/µl 1	Exactive pmol/µl 2	Exactive pmol/µl 3	PREST 1 pmol	PREST 2 pmol	PREST 3 pmol	Proteina 1 pmol	Proteina 2 pmol	Proteina 3 pmol
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	5,535328	3,057585	4,098071
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	3,962545	2,333958	3,207972
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	6,236442	3,259643	3,993737
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	#VALUE!	#VALUE!	#VALUE!
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	4,378646	2,793736	#VALUE!
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	#VALUE!	#VALUE!	#VALUE!
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	#VALUE!	2,276252	#VALUE!
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	4,733855	2,682069	3,498447
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	5,191679	3,568942	3,497855
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	4,943879	2,870167	3,649909
0,121	0,0495	232,98455	140,682	59,879442	28,19113	17,02252	2,964032	4,934857	3,171807	3,21805
0,004	0,0325	130,83738	149,56	50,714554	0,52335	0,59824	1,648223	1,828636	0,854705	2,083848
0,004	0,0325	130,83738	149,56	50,714554	0,52335	0,59824	1,648223	#VALUE!	1,122896	2,358607
0,004	0,0325	130,83738	149,56	50,714554	0,52335	0,59824	1,648223	1,694344	#VALUE!	#VALUE!
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	#VALUE!	#VALUE!	2,065223
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	0,340329	0,225076	0,258617
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	0,297857	0,200712	0,23486
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	0,278931	0,198747	0,159273
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	0,347426	0,254616	0,260099
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	0,297568	0,236095	0,251479
0,005	0,0032	160,39251	113,243	49,792301	0,801963	0,566215	0,159335	0,32619	#VALUE!	#VALUE!
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	7,415865	5,331402	5,874169
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	6,791159	#VALUE!	#VALUE!
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	6,679354	4,985963	5,473414
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	7,681211	4,672469	6,761402
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	8,992427	#VALUE!	#VALUE!
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	6,572693	4,066979	6,069724
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	6,234918	4,655659	4,813882
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	5,888459	4,52473	4,704318
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	7,039473	5,947942	6,267423
0,066	0,1232	127,7533	101,472	43,487725	8,431718	6,697152	5,357688	5,98711	4,317721	4,395876

Tabla complementaria 2 continuación

Nombres de proteínas	Nombre del gen	Nombre del Secuencia	Relación de H/L de mezcla maestra		Relación de H/L de mezcla maestra		Relación de H/L de mezcla CV (%)
			(1)	(2)	(2)	(3)	
Endonucleasa Flap 1	FEN1	AVDLIQK	0,72265	0,74027	19,67	21,43	1,2719
Endonucleasa Flap 1	FEN1	EAHQLFLEPEVLDPESVELK	0,85331	0,83635			10,53
Endonucleasa Flap 1	FEN1	HLTASEAK	0,69703	0,63769			1,3723
Endonucleasa Flap 1	FEN1	LDPNKYPVPENWLHK	0,73377	0,68			1,0235
Endonucleasa Flap 1	FEN1	LPIQEFHLSR	0,65281	0,4747			1,2048
Endonucleasa Flap 1	FEN1	SIEEIVR	1,1354				1,0466
Endonucleasa Flap 1	FEN1	YAAATEDMDCLTFGSPVLMR	0,73515	0,45581	NaN	NaN	1,1345
Endonucleasa Flap 1	FEN1	YVPENWLHK	0,72303	0,70528			1,1416
Oncogén celular fos;G	FOS	ELTDTLQAETDQLEDEK	NaN	NaN	-	26,11	NaN
Oncogén celular fos;G	FOS	GKVEQLSPEEEEEK	NaN	0,0090707			0,050604
Oncogén celular fos;G	FOS	SALQTEIANLLK	0,015533	0,1013178			0,054066
Oncogén celular fos;G	FOS	VEQLSPEEEEEK	NaN	NaN			NaN
Proteína de choque térmico de 70 kDa«	HSPA4	EDQYDHLDAADMTK	0,2685	0,20351	78,77	7,31	0,79843
Proteína de choque térmico de 70 kDa	HSPA4	LNLQNK	0,25909	0,21666			0,94161
Proteína de choque térmico de 70 kDa	HSPA4	NKEDQYDHLDAADMTK	0,21899	0,21981			0,84382
Proteína de choque térmico de 70 kDa	HSPA4	QIQQYMK	0,95485	NaN			NaN
Proteína de choque térmico de 70 kDa	HSPA4	QSLTMDPWK	0,24663	0,22089			0,89363
Proteína de choque térmico de 70 kDa	HSPA4	STNEAMIEVMNKK	0,25098	0,24797			1,1176
Li activador de GTPasa Ras	IQGAP1	ILAIGLINEALDEGDAQK	0,26959	0,27248	9,98	5,48	1,4819
Li activador de GTPasa Ras	IQGAP1	LEGVLAEVAQHYQDTLIR	0,27918	0,273			1,4224
Li activador de GTPasa Ras	IQGAP1	QLSSSVTGLTNIIEENCQR	0,2718	0,23809			1,4471
Li activador de GTPasa Ras	IQGAP1	TLQALQIPAAK	0,218	0,25878			1,5677
Li activador de GTPasa Ras	IQGAP1	YLDELIMK	0,27889	0,26493			1,3441
Proteína activada por mitógeno	MAP2K1IP1	ELAPLFEELR	0,32694	0,056059	-	-	1,6953
Proteína activada por mitógeno	MAP2K1IP1	KLPSVEGLHAIWSDR	NaN	NaN			1,1365
Proteína activada por mitógeno	MAP2K1IP1	VANDNAPEHALRPGFLSTFALATI	NaN	NaN			0,89538
Dominio de quinasa de linaje mixto	MLKL	APVAIK	0,39067	NaN	-	-	1,0384
proteína ribosómica 39S L	MRPL50	AYTPPEDLQSR	0,40095	NaN	2,84	5,03	2,6456
proteína ribosómica 39S L	MRPL50	EKEPVVVVETVEEK	0,41737	NaN			2,3002
proteína ribosómica 39S L	MRPE50	LESYVK	NaN	NaN			NaN

Tabla complementaria 2 continuación

µl PRESTs	(1+2)	µl PRESTs	(3)	Exactive	Exactive	Exactive	Exactive	pmol/µl 1	pmol/µl 2	Exactive	pmol/µl 3	pmol	PrEST 1	pmol	PrEST 2	pmol	PrEST 3	pmol	Proteína 1	pmol	Proteína 2	pmol	Proteína 3
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	3,921954	3,663829	2,893179												
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	4,63107	3,121558													
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	3,78291	4,13936													
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	3,982304	2,328146													
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	3,54292	2,740547													
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	6,162024	2,380691													
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	#VALUE!	#VALUE!													
0.035	0.061	155,06238	141,409	37,29001	5,427183	4,949315	2,274691	2,274691	3,989794	2,580637													
0.007	0.0033	151,55943	141,201	52,196632	1,060916	0,988407	0,172249	0,172249	#VALUE!	#VALUE!													
0.007	0.0033	151,55943	141,201	52,196632	1,060916	0,988407	0,172249	0,172249	#VALUE!	#VALUE!													
0.007	0.0033	151,55943	141,201	52,196632	1,060916	0,988407	0,172249	0,172249	0,008966	0,008716													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	0,013025	0,009313													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	#VALUE!	#VALUE!													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	3,728623	2,440554													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	3,597947	2,878211													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	3,041084	2,438997													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	13,25987	2,474458													
0.073	0.0447	190,231	154,209	68,382353	13,88686	11,25726	3,056691	3,056691	3,424917	#VALUE!													
0.109	0.054	98,606097	72,632	27,555556	10,74806	7,916888	1,488	1,488	3,485325	2,731551													
0.109	0.054	98,606097	72,632	27,555556	10,74806	7,916888	1,488	1,488	2,897571	2,205067													
0.109	0.054	98,606097	72,632	27,555556	10,74806	7,916888	1,488	1,488	2,98929	2,16531													
0.109	0.054	98,606097	72,632	27,555556	10,74806	7,916888	1,488	1,488	3,000645	2,153285													
0.109	0.054	98,606097	72,632	27,555556	10,74806	7,916888	1,488	1,488	2,343078	2,048732													
0.006	0.0033	149,10541	128,996	49,884398	0,894632	0,773976	0,164619	0,164619	2,997528	2,097421													
0.006	0.0033	149,10541	128,996	49,884398	0,894632	0,773976	0,164619	0,164619	0,292491	0,043388													
0.005	0.0048	105,25905	95,876	38,828103	0,526295	0,47938	0,186375	0,186375	#VALUE!	#VALUE!													
0.006	0.0042	121,96997	107,299	33,817122	0,73182	0,643794	0,142032	0,142032	0,205608	0,166876													
0.006	0.0042	121,96997	107,299	33,817122	0,73182	0,643794	0,142032	0,142032	0,293423	0,147486													
0.006	0.0042	121,96997	107,299	33,817122	0,73182	0,643794	0,142032	0,142032	0,30544	0,37576													
0.006	0.0042	121,96997	107,299	33,817122	0,73182	0,643794	0,142032	0,142032	#VALUE!	0,435675													
0.006	0.0042	121,96997	107,299	33,817122	0,73182	0,643794	0,142032	0,142032	#VALUE!	0,405725													

Tabla complementaria 2 continuación

Nombres de proteínas	Nombre del gen Secuencia	Relación de H/L de mezcla maestra		Relación de H/L de mezcla maestra		CV (%)
		(1)	(2)	(2)	(3)	
Proteína ribosómica 28S S	MRPS23	0,33211	0,30866	51,08	2,8825	3,38
Proteína ribosómica 28S S	MRPS23	0,31151	0,33385		2,7777	
Proteína ribosómica 28S S	MRPS23	0,36147	0,061638		2,6945	
Proteína ribosómica 28S S	MRPS23	0,30384	0,3357		NaN	
Proteína ribosómica 28S S	MRPS28	0,79424	0,59711	14,13	NaN	11,34
Proteína ribosómica 28S S	MRPS28	0,48648	0,62859		1,2025	
Proteína ribosómica 28S 5	MRPS28	0,87278	0,77368		1,4122	
Purina nucleósido fosf	NP	0,65455	0,63682	21,02	0,89372	8,08
Purina nucleósido fosf	NP	0,57033	0,47254		0,98592	
Purina nucleósido fosf	NP	0,56081	0,52235		0,98611	
Purina nucleósido fosf	NP	0,58408	0,57218		0,85084	
Purina nucleósido fosf	NP	NaN	0,50154		0,84573	
Purina nucleósido fosf	NP	0,26064	0,27356		0,79355	
Purina nucleósido fosf	NP	0,54236	0,51975		0,88994	
Purina nucleósido fosf	NP	0,68813	0,51622			
Purina nucleósido fosf	NP	0,18498	0,18395	6,50		
Poli [ADP-ribosa] polim	PARP4	NaN	NaN		NaN	-
Poli [ADP-ribosa] polim	PARP4	NaN	NaN		NaN	
Poli [ADP-ribosa] polim	PARP4	1089268	0,16778		NaN	
Subunidad de prefoldina 1	PFDN1	2,89	NaN	12,33	NaN	8,38
Subunidad de prefoldina 1	PFDN1	0,17447	0,1967		1,8144	
Subunidad de prefoldina 1	PFDN1	0,18176	NaN		NaN	
Subunidad de prefoldina 1	PFDN1	NaN	0,16018		1,6113	
Subunidad de prefoldina 1	PFDN1	0,47962	0,20263	7,16	NaN	
Peptidil-propilil-cis-trans	PPIB	0,50189	0,47562		1,0739	7,51
Peptidil-propilil-cis-trans	PPIB	0,5755	0,51014		1,2206	
Peptidil-propilil-cis-trans	PPIB	0,46154	0,453		1,3142	
Peptidil-propilil-cis-trans	PPIB	0,49743	0,49432		1,1893	
Peptidil-propilil-cis-trans	PPIB	0,48337	0,49474		1,2376	
Peptidil-propilil-cis-trans	PPIB		0,49474		1,3259	

Tabla complementaria 2 continuación

µl PrESTs (1+2)	0,006	0,0024	198,09216	Exactive pmol/µl 1	Exactive pmol/µl 2	Exactive pmol/µl 3	pmol PrEST 1	pmol PrEST 2	pmol PrEST 3	pmol Proteína 1	pmol Proteína 2	pmol Proteína 3
0,006	0,0024	198,09216	181,075	69,643134	1,188553	1,08645	0,167144	0,39473	0,335344	0,39473	0,335344	0,481791
0,006	0,0024	198,09216	181,075	69,643134	1,188553	1,08645	0,167144	0,370246	0,362711	0,370246	0,362711	0,464275
0,006	0,0024	198,09216	181,075	69,643134	1,188553	1,08645	0,167144	0,429626	0,066967	0,429626	0,066967	0,450368
0,006	0,0024	198,09216	181,075	69,643134	1,188553	1,08645	0,167144	0,36113	0,364721	0,36113	0,364721	#VALUE!
0,005	0,0081	191,91632	145,872	66,25803	0,959582	0,72936	0,53669	0,762138	0,435508	0,762138	0,435508	#VALUE!
0,005	0,0081	191,91632	145,872	66,25803	0,959582	0,72936	0,53669	0,466817	0,458468	0,466817	0,458468	0,64537
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	4,04012	2,796299	4,04012	2,796299	2,593789
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	3,520284	2,07494	3,520284	2,07494	2,861376
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	3,461523	2,425388	3,461523	2,425388	2,861927
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	#VALUE!	2,512462	#VALUE!	2,512462	2,469341
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	1,201212	2,20228	1,201212	2,20228	2,454511
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	1,608765	1,201212	1,608765	1,201212	#VALUE!
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	3,347643	2,28224	3,347643	2,28224	2,303072
0,005	0,0084	112,22478	79,837	33,590734	6,172363	4,391035	2,902239	4,247388	2,26674	4,247388	2,26674	2,582819
0,005	0,0037	144,98538	124,938	55,100415	0,724927	0,62469	0,203872	0,134097	0,114912	0,134097	0,114912	#VALUE!
0,005	0,0037	144,98538	124,938	55,100415	0,724927	0,62469	0,203872	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
0,005	0,0037	144,98538	124,938	55,100415	0,724927	0,62469	0,203872	0,064713	0,10481	0,064713	0,10481	#VALUE!
0,023	0,0071	193,04964	191,964	33,110306	4,440142	4,415172	0,235083	#VALUE!	#VALUE!	#VALUE!	#VALUE!	0,428535
0,023	0,0071	193,04964	191,964	33,110306	4,440142	4,415172	0,235083	0,774672	0,868464	0,774672	0,868464	#VALUE!
0,023	0,0071	193,04964	191,964	33,110306	4,440142	4,415172	0,235083	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
0,023	0,0071	193,04964	191,964	33,110306	4,440142	4,415172	0,235083	0,80704	0,707222	0,80704	0,707222	0,37879
0,023	0,0071	193,04964	191,964	33,110306	4,440142	4,415172	0,235083	#VALUE!	0,894646	#VALUE!	0,894646	#VALUE!
0,531	0,6195	99,552439	58,765	22,904459	52,86234	31,20422	14,18931	25,35384	13,03213	25,35384	13,03213	15,2379
0,531	0,6195	99,552439	58,765	22,904459	52,86234	31,20422	14,18931	26,53108	14,84135	26,53108	14,84135	17,31947
0,531	0,6195	99,552439	58,765	22,904459	52,86234	31,20422	14,18931	30,42228	15,91852	30,42228	15,91852	18,64759
0,531	0,6195	99,552439	58,765	22,904459	52,86234	31,20422	14,18931	24,39809	14,13551	24,39809	14,13551	16,87535
0,531	0,6195	99,552439	58,765	22,904459	52,86234	31,20422	14,18931	26,29532	15,42487	26,29532	15,42487	17,56069
0,531	0,6195	99,552439	58,765	22,904459	52,86234	31,20422	14,18931	25,55207	15,43797	25,55207	15,43797	18,81361

Tabla complementaria 2 continuación

Nombres de proteínas	Nombre del gen	Secuencia	Relación de H/L de mezcla maestra (1)		Relación de H/L de mezcla maestra (2)		Relación de H/L de mezcla maestra (3)	
			CV (%)	Relación de H/L de mezcla maestra	CV (%)	Relación de H/L de mezcla maestra	CV (%)	Relación de H/L de mezcla maestra
Peroxirredoxina 6	PRDX6	DFTVCTTELGR	0,70188	6,96	0,68855	7,92	0,98276	9,15
Peroxirredoxina 6	PRDX6	DINAYNCEEPTEK	0,75055		0,74788		1,1286	
Peroxirredoxina 6	PRDX6	ELAILLGMLDPAEK	0,77082		0,61275		1,2449	
Peroxirredoxina 6	PRDX6	ELAILLGMLDPAEKDEK	0,62879		0,65579		1,2636	
Peroxirredoxina 6	PRDX6	FHDFLGDSWGILFSHPR	NaN		0,78925		NaN	
Peroxirredoxina 6	PRDX6	GMPVTAR	0,71311		0,79113		1,0751	
Peroxirredoxina 6	PRDX6	LAPEFAK	0,73135		0,65758		1,1849	
Peroxirredoxina 6	PRDX6	LIALSIDSVEDHLAWSK	0,72068		0,73169		1,1426	
Peroxirredoxina 6	PRDX6	LFPPIIDDR	0,74256		0,74049		1,1585	
Peroxirredoxina 6	PRDX6	VVVFVFGPDK	0,6815		0,74577		1,1822	
Peroxirredoxina 6	PRDX6	VVVFVFGPDKK	0,81434		0,72122		0,94586	
proteasa 26S reguladora	PSMC3	AVCVEAGMIALR	0,89983	24,43	0,51761	10,16	NaN	-
proteasa 26S reguladora	PSMC3	GATELTHEDYMEGILEVQAK	NaN		NaN		NaN	
proteasa 26S reguladora	PSMC3	IMQIHSR	0,58166		0,60039		1,0965	
proteasa 26S reguladora	PSMC3	MNVSPDVNYEELAR	0,62842		0,63254		NaN	
proteína sigma 14-3-3	SFN	IIDSAR	NaN		NaN		NaN	16,05
proteína sigma 14-3-3	SFN	SAYQEAMDISK	0,41269		0,41814		1,1131	
proteína sigma 14-3-3	SFN	SNEEGSEKGPVPR	0,3446		0,35575		0,75982	
proteína sigma 14-3-3	SFN	VETELQGVGDTVLGLLDSHLIK	NaN		0,43773		NaN	
proteína sigma 14-3-3	SFN	VLSSIEQK	0,39497		0,35675		0,87046	
proteína sigma 14-3-3	SFN	YLAEVATGDDK	0,41281		0,38665		0,98489	
proteína sigma 14-3-3	SFN	YLAEVATGDDKK	0,404		0,41942		0,79403	
Subunidad S de complejo FACT	SSRP1	ADVIGATGDAICIFR	0,85761	3,66	0,73149	6,06	1,1277	5,26
Subunidad S de complejo FACT	SSRP1	ELQCLTPR	0,90583		0,73695		1,2494	
Subunidad S de complejo FACT	SSRP1	IPYTTVLR	0,84574		0,79479		1,2167	
Subunidad S de complejo FACT	SSRP1	LFLPHK	NaN		0,68587		NaN	
Subunidad 1 de complejo THO	THOC1	AVNNSNYGWR	0,43053	26,28	0,34141	24,03	NaN	43,19
Subunidad 1 de complejo THO	THOC1	LWNLCPDNMEACK	0,55515		0,38786		2,8589	
Subunidad 1 de complejo THO	THOC1	SLPEYLENMVIK	0,50412		0,38951		2,9525	
Subunidad 1 de complejo THO	THOC1	TGEDEDEEDNDALLK	0,28678		0,21856		1,1674	
Nucleoproteína TPR	TPR	ILLSQTTGVAIPLHASSLDDVSLAS	0,1447	7,15	0,16851	10,98	NaN	16,59
Nucleoproteína TPR	TPR	ITELQLK	0,17454		0,16679		1,6804	

Tabla complementaria 2 continuación

µl PrESTs (1+2)	µl PrESTs (3)	Exactive pmol/µl 1	Exactive pmol/µl 2	Exactive pmol/µl 3	pmol PrEST 1	pmol PrEST 2	pmol PrEST 3	pmol Proteína 1	pmol Proteína 2	pmol Proteína 3
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	14,27057	13,1034	12,45393
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	15,26013	14,23247	14,30207
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	15,67226	11,66089	15,77587
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	12,78451	12,47896	16,01285
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	#VALUE!	15,01976	#VALUE!
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	14,4989	15,05554	13,6241
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	14,86976	12,5140 2	15,01553
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	14,65281	13,92437	14,47949
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	15,09768	14,09184	14,68098
0,202	0,3388	100,65311	94,21	37,403786	20,33193	19,03042	12,6724	13,85621	14,19232	14,98131
0,102	0,0294	167,07191	156,456	61,444877	2,840222	2,659752	1,806479	16,5571	13,72512	11,98632
0,102	0,0294	167,07191	156,456	61,444877	2,840222	2,659752	1,806479	2,555717	1,376714	#VALUE!
0,102	0,0294	167,07191	156,456	61,444877	2,840222	2,659752	1,806479	#VALUE!	#VALUE!	#VALUE!
0,102	0,0294	167,07191	156,456	61,444877	2,840222	2,659752	1,806479	1,784853	1,596889	#VALUE!
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	#VALUE!	#VALUE!	#VALUE!
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	3,99777	3,217858	3,407697
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	3,338175	2,737727	2,326149
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	#VALUE!	3,368616	#VALUE!
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	3,826114	2,745422	2,664867
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	3,998932	2,975522	3,015189
0,096	0,102	100,90731	80,163	30,014189	9,687101	7,695648	3,061447	3,913589	3,227709	2,430881
0,017	0,0343	128,21107	139,714	48,156415	2,179588	2,375138	1,651765	1,869237	1,73739	1,862695
0,017	0,0343	128,21107	139,714	48,156415	2,179588	2,375138	1,651765	1,974336	1,750358	2,063715
0,017	0,0343	128,21107	139,714	48,156415	2,179588	2,375138	1,651765	1,843365	1,887736	2,009703
0,017	0,0343	128,21107	139,714	48,156415	2,179588	2,375138	1,651765	#VALUE!	1,629036	#VALUE!
0,006	0,0036	133,17121	131,715	33,493744	0,799027	0,79029	0,120577	0,344005	0,269813	#VALUE!
0,006	0,0036	133,17121	131,715	33,493744	0,799027	0,79029	0,120577	0,44358	0,306522	0,344719
0,006	0,0036	133,17121	131,715	33,493744	0,799027	0,79029	0,120577	0,402806	0,307826	0,356005
0,006	0,0036	133,17121	131,715	33,493744	0,799027	0,79029	0,120577	0,229145	0,172726	0,140762
0,021	0,0086	186,99624	129,566	51,550628	3,926921	2,720886	0,443335	0,568225	0,458496	#VALUE!
0,021	0,0086	186,99624	129,566	51,550628	3,926921	2,720886	0,443335	0,685405	0,453817	0,744981

Tabla complementaria 2 continuación

Nombres de proteínas	Nombre del gen	Secuencia	Relación de H/L de mezcla maestra (1)	CV (%)	Relación de H/L de mezcla maestra (2)	CV (%)	Relación de H/L de mezcla maestra (3)	CV (%)
Nucleoproteína TPR	TPR	LESALTELEQLR	0,1666		0,15943		1,1901	
Nucleoproteína TPR	TPR	LESALTELEQLRK	0,17493		0,17872		1,3357	
Nucleoproteína TPR	TPR	NIEELQQNQNR	0,17718		0,18774		1,3128	
Nucleoproteína TPR	TPR	QHMQQLVDSIVR	0,16848		0,21389		1,7312	
Complejo de citocromo b-c1	UQCRC1	ADLTEYLSTHYK	1,6066	5,68	1,5721	5,68	1,3068	12,57
Complejo de citocromo b-c1	UQCRC1	DFWNYLHATAFQGTPLAGAVEG	1,524		1,4183		1,3969	
Complejo de citocromo b-c1	UQCRC1	MVLAAGGVHEHQQLDLAQK	1,707		1,5624		1,6586	
Endoplasmina de transición	VCP	DHFEFAMR	0,20879	29,20	0,20611	21,24	1,4096	26,47
Endoplasmina de transición	VCP	ESIESEIR	0,075211		0,16428		NaN	
Endoplasmina de transición	VCP	GFGSFR	0,21737		0,22985		1,287	
Endoplasmina de transición	VCP	KYEMFAQTLQQR	0,12676		NaN		0,47634	
Endoplasmina de transición	VCP	MTNGFSGADLTEICQR	0,23492		0,2695		1,5438	
Endoplasmina de transición	VCP	QTNPSAMEVEEODPVPEIR	0,14391		0,15975		1,292	
Endoplasmina de transición	VCP	RDFEEAMR	0,20524		0,17819		1,459	
Endoplasmina de transición	VCP	SVSDNDR	0,19148		0,15796		1,2779	
Endoplasmina de transición	VCP	YEMFAQTLQQR	0,20698		0,24338		1,4679	
Vimentina	VIM	DNLAEIMR	0,813	15,32	0,79552	16,24	1,6492	33,78
Vimentina	VIM	EEAENTLQSF	0,78127		0,81105		1,4753	
Vimentina	VIM	EKLQEEMLQR	NaN		NaN		NaN	
Vimentina	VIM	ILLAELEQLK	0,7214		0,71178		1,3289	
Vimentina	VIM	ILLAELEQLKGGK	0,7299		NaN		2,7399	
Vimentina	VIM	LGDLYEEEMR	0,52115		0,45107		0,90103	
Vimentina	VIM	LQEEMLQR	0,87912		0,8627		1,8464	
Vimentina	VIM	QDVDNASLAR	0,82023		0,89058		1,652	
Vimentina	VIM	QVDQLTNDK	0,79102		0,72609		1,3024	
Vimentina	VIM	RQVDQLTNDK	0,84806		0,70013		1,2391	
Vimentina	VIM	VEVERDNLAEIMR	0,58525		0,77955		1,0746	
Homo letal para hembras (2) D	WTAP	EGNTTEDDFPSSPGNGNK	NaN	-	NaN	-	NaN	-
Homo letal para hembras (2) D	WTAP	LTINGPSNGSSSR	NaN		NaN		NaN	
Homo letal para hembras (2) D	WTAP	QQLAQYQQQQSQASAPSTR	0,05126		NaN		0,67594	
Homo letal para hembras (2) D	WTAP	TSGSGFHR	NaN		NaN		NaN	

CV mediano 12,33  
CV promedio 18,39

Tabla complementaria 2 continuación

µl PRESTs (1+2)	µl PRESTs (3)	Exactive pmol/µl 1	Exactive pmol/µl 2	Exactive pmol/µl 3	pmol PREST 1	pmol PREST 2	pmol PREST 3	pmol Proteína 1	pmol Proteína 2	pmol Proteína 3
0,021	0,0086	186,99624	129,566	51,550628	3,926921	2,720886	0,443335	0,654225	0,433791	0,527613
0,021	0,0086	186,99624	129,566	51,550628	3,926921	2,720886	0,443335	0,686936	0,486277	0,592163
0,021	0,0086	186,99624	129,566	51,550628	3,926921	2,720886	0,443335	0,695772	0,510819	0,582011
0,021	0,0086	186,99624	129,566	51,550628	3,926921	2,720886	0,443335	0,661608	0,58197	0,767502
0,006	0,0215	179,22945	128,578	56,812437	1,075377	0,771468	1,221467	1,7277	1,212825	1,596214
0,006	0,0215	179,22945	128,578	56,812437	1,075377	0,771468	1,221467	1,638874	1,094173	1,706268
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	1,835668	1,205342	2,025926
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	4,606713	4,212569	5,038285
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	1,659445	3,357629	#VALUE!
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	2,796814	4,697778	4,60008
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	5,183241	5,508163	1,702566
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	3,17521	3,265043	5,517951
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	4,528386	3,641928	4,617951
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	4,224787	3,228458	5,214854
0,099	0,0421	222,86724	206,449	84,899424	22,06386	20,43845	3,574266	4,566777	4,97431	4,567554
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	39,41833	29,41602	5,246665
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	37,8799	29,99027	44,7033
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	#VALUE!	29,99027	39,98956
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	#VALUE!	28,8303	#VALUE!
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	34,9771	26,31956	36,02123
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	35,38922	#VALUE!	74,26788
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	25,26797	16,67926	24,42337
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	42,62416	31,90014	50,04862
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	39,76887	32,93106	44,7792
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	38,35263	26,8487	35,30292
0,297	0,4921	163,24925	124,502	55,082409	48,48503	36,97709	27,10605	41,11821	25,88877	33,58711
0,0028	0,0028	178,53242	165,579	59,162252	1,071195	0,993474	0,165654	#VALUE!	#VALUE!	29,12817
0,006	0,0028	178,53242	165,579	59,162252	1,071195	0,993474	0,165654	#VALUE!	#VALUE!	#VALUE!
0,006	0,0028	178,53242	165,579	59,162252	1,071195	0,993474	0,165654	0,054909	#VALUE!	0,111972
0,006	0,0028	178,53242	165,579	59,162252	1,071195	0,993474	0,165654	#VALUE!	#VALUE!	#VALUE!

Para evaluar de forma independiente la precisión de esta etapa de cuantificación de proteínas absoluta, se compararon las relaciones determinadas a partir de “péptidos trípticos límite” (los que no tenían arg o lys) con las determinadas a partir de las versiones más largas del péptido que contiene uno o dos sitios de escisión trípticos ausentes. Estos péptidos son muy problemáticos para métodos basados en patrón peptídico tales como AQUA, pero en las mediciones de los inventores se midieron relaciones muy similares para dichos péptidos. Esto muestra que la digestión continúa de forma idéntica para PrEST y proteína endógena (Tabla 1). Por lo tanto, lejos de introducir incertidumbre, en el enfoque de SILAC - PrEST estos péptidos pueden proporcionar información de cuantificación adicional.

10 **Tabla 1: Comparación de péptidos trípticos límite y péptidos con los sitios de escisión tríptica ausentes.** Se representan péptidos con una o dos escisiones ausentes así como sus relaciones. Las relaciones de las dos versiones varían en promedio en 19 %, lo que está en el intervalo normal de variación de péptidos derivados de una proteína.

Nombres deSecuencia genes		Escisiones ausentes	Relación H/L de mezcla maestra 1	CV (%)	Relación H/L de mezcla maestra 2	CV (%)	Relación H/L de mezcla maestra 3	CV (%)
HSPA4	ELAILLGMLDPAEK	0	0,707	16,2	0,803	6,7	1,072	19,7
HSPA4	ELAILLGMLDPAEKDEK	1	0,562		0,730		1,418	
HSPA4	EDQYDHLDAADMKT	0	0,220	9,2	0,132	25,8	0,558	35,6
HSPA4	NKEDQYDHLDAADMTK	1	0,193		0,191		0,933	
ATP5B	VKDSGAPIK	0	0,738	11,8	0,452	46,9	1,009	6,5
ATP5B	VLD SGAPIKIPVGPETLGR	1	0,872		0,901		1,107	
PPIB	DKPLKDVIADCGK	2	0,526	6,9	0,436	14,6	0,889	23,2
PPIB	DVIIADCGK	0	0,580		0,354		1,237	
FASN	QQEQQVPILEK	0			0,627	18,6	1,040	13,9
FASN	RQQEQQVPILEK	1			0,481		0,853	
FEN1	LDPNKYPVPENWLHK	1	0,680	8,1	0,632	0,3	1,279	14,4
FEN1	YPVPENWLHK	0	0,607		0,629		1,043	
SFN	EKVETELQGVC DTVLGLLDSHLI K	1	0,442	6,3	0,389	6,4	1,188	1,2
SFN	VETELQGVC DTVLGLLDSHLIK	0	0,483		0,426		1,168	
SFN	SNEEGSEEK	0	0,286	11,0				
SFN	SNEEGSEEKGPEVR	1	0,334					
SFN	YLAEVATGDDK	0	0,371	1,3	0,395	19,7	1,012	8,5
SFN	YLAEVATGDDKK	1	0,364		0,299		1,142	
TPR	LESALTELEQLR	0	0,139	17,4	0,121	13,0	1,243	10,9
TPR	LESALTELEQLRK	1	0,177		0,145		1,064	
VCP	DHFEEARMR	0	0,187	11,1	0,134	35,7	1,712	59,7
VCP	RDHFEEAMR	1	0,218		0,224		0,696	
VCP	YEMFAQTLQQSR	0	0,169	4,4	0,146	6,6	1,584	37,1
VCP	KYEMFAQTLQQSR	1	0,159		0,133		0,926	
VIM	QVDQLTNDK	0	0,620	8,50	0,813	2,3		
VIM	RQVDQLTNDK	1	0,699		0,788			
VIM	LQEMLQR	0	0,868	3,9	0,834	65,0		
VIM	EKLQEMLQR	1	0,821		0,309			

15 Para evaluar el grado de variabilidad asociado con ambas etapas del procedimiento de cuantificación absoluta, se repitió el flujo de trabajo completo dos veces más, incluyendo cuantificación de PrEST y generación de mezcla maestra así como medición de abundancia celular de las proteínas diana. Este análisis mostró que los errores típicos de la media asociados con todas las etapas juntas son en promedio 24 %. Este valor es excelente y hasta donde alcanza el conocimiento de los inventores la determinación más precisa de niveles de expresión celular indicados hasta la fecha. Lo que es aún más importante, los errores de cada una de las etapas en el flujo de trabajo para cada una de las proteínas son inmediatamente aparentes a partir de los CV individuales. Por lo tanto todas las mediciones del nivel de expresión de proteínas pueden clasificarse y aceptarse o descartarse de acuerdo con la confianza de las mediciones. La Figura 4 presenta ejemplos típicos de determinación de expresión de proteínas a partir de las mediciones por triplicado. La comparación de las distribuciones de relación de péptidos con respecto a la variabilidad de los valores de proteína medios reveló que la preparación de la mezcla maestra contribuía a la mayor variabilidad mientras que los errores debidos a determinación de relación de SILAC fueron algo menores. La preparación automática de la mezcla maestra podría conducir por lo tanto a mejoras adicionales en el futuro.

30 **Determinación del número de copias de proteínas en células HeLa** – A continuación se usaron los valores absolutos para cantidades de proteínas en el lisado de células HeLa de los inventores para calcular los números de copias correspondientes en células. Los números de células HeLa se determinaron automáticamente en un contador celular (véase Procedimientos Experimentales). Dada la cantidad conocida de cada PrEST y sus relaciones de SILAC con

respecto a las proteínas endógenas se determinaron los números de copias celulares de 37 proteínas diferentes. Se consiguió una precisión muy alta de cuantificación absoluta hasta una distancia de un error típico del 25 % para 30 de 37 proteínas (Tabla 2).

5 **Tabla 2: números de copias de proteínas por célula HeLa**

Nombres de proteínas	Nombres de genes	Mediana	ETM (%)*	Mezcla maestra 1	Mezcla maestra 2	Mezcla maestra 3
Proteína sigma 14-3-3	SFN	2.104.742	9,57	2.128.717	1.562.390	2.104.742
Subunidad reguladora de proteasa 26S 6A	PSMC3	1.009.040	6,71	1.009.040	985.010	1.211.206
Proteína ribosómica 28S S23, mitocondrial	MRPS23	202.529	19,64	202.529	161.109	308.977
Proteína ribosómica 28S S35, mitocondrial	MRPS28	516.278	15,46	586.618	337.285	516.278
Proteína ribosómica 39S L50, mitocondrial	MRPL50	212.893	20,00	170.320	255.465	
Proteína de tipo AFG3 2	AFG3L2	335.545	20,37	335.545	363.149	173.343
Subunidad beta de ATP sintasa, mitocondrial	ATP5B	4.870.803	5,63	5.431.604	4.870.803	4.476.459
Carbonil reductasa [NADPH] 3	CBR3	101.019	63,47	101.019	60.715	498.397
Proteína de cuerpo multivesicular con carga 6	CHMP6	133.137	24,10	154.946	61.839	133.137
Subunidad 5 de complejo de signalosoma COP9	COPS5	287.189	13,59	287.189	211.517	343.078
Citocromo b5 reductasa 4	CYB5R4	10.537	**	10.537		
Subunidad 1 de complejo de citocromo b-c1, mitocondrial	UQCRC1	1.032.315	8,96	1.032.315	808.601	1.099.145
Acil coenzima A tioéster hidrolasa citosólica	ACOT7	455.871	0,46	457.985	453.757	
Proteína asociada a balsa lipídica de retículo endoplásmico 2	ERLIN2	135.785	18,00	218.008	127.563	135.785
Enoil-CoA hidratasa, mitocondrial	ECHS1	2.162.058	15,20	2.705.599	1.574.948	2.162.058
Factor de inicio de la traducción eucariota 3 subunidad 6	EIF3E	1.298.361	9,85	1.009.294	1.298.361	1.422.955
Subunidad SSPR1 del complejo de FACT	SSRP1	1.054.400	4,43	1.086.956	937.213	1.054.400
Ácido graso sintasa	FASN	3.361.337	13,11	4.093.238	2.575.577	3.361.337
Endonucleasa Flap1	FEN1	2.215.232	6,87	2.215.232	2.220.140	1.789.805
Proteína de choque térmico de 70 kDa 4	HSPA4	1.719.164	11,22	1.855.515	1.258.240	1.719.164
Antígeno asociado a carcinoma hepatocelular 59	C9orf78	140.949	19,08	1.577.757	140.949	114.844
Proteína de armazón de proteína quinasa activada por mitógeno 1	MAP2K1IP1	160.325	10,12	160,25	205.536	148.606
Nucleoproteína TPR	TPR	306.362	14,08	343.837	208.601	306.362
Peptidil-prolil cis-trans isomerasa B	PPIB	11.155.435	17,40	14.871.040	8.035.119	11.155.435
Peroxisredoxina 6	RPDX6	8.815.042	3,13	9.010.737	8.118.496	8.815.042
Subunidad de prefoldina 1	PFDN1	358.511	22,03	383.937	358.511	171.199
Regulador de corte y empalme pre ARNm WTAP	WTAP	72.1999	-	72.199	-	-
ARN helicasa dependiente de ATP probable DDX20	DDX20	268.121	-	268.121	-	-
Protooncogén c-Fos	FOS	3.651	15,95	4.559	2.575	3.651
Purina nucleósido fosforilasa	NP	1.618.680	10,11	1.832.987	1.284.556	1.618.680
Proteína de tipo activadora de GTPasa Ras IQGAP1	IQGAP1	1.322.762	15,42	1.667.348	963.468	1.322.762
Proteína de unión a ARN de interacción tallo-bucle SRA, mitocondrial	C14orf156	1.482.399	15,25	1.557.983	919.914	1.482.399
Subunidad beta de proteína de complejo T 1	CCT2	4.352.706	29,38	8.283.044	3.162.779	4.352.706
Subunidad 1 de complejo THO	TH0C1	191.319	9,53	211.743	151.702	191.319
ATPasa de retículo endoplásmico transicional	VCP	2.343.243	10,89	2.343.2423	1.716.701	2.493.783
Vimentina	VIM	20.600.599	8,73	20.600.599	17.557.991	23.805.318
Proteína de dedo de zinc 828	C13orf8	167.150	29,45	117.929	216.371	

\*Error típico de la media (ETM) para las tres repeticiones en porcentaje.

\*\*sin datos válidos obtenidos

10 Los números de copias celulares se conocen solamente para muy pocas proteínas y por lo tanto es interesante relacionar estos números de copias con las funciones conocidas de las proteínas (Tabla Complementaria 3). La proteína citoesquelética vimentina forma filamentos intermedios y fue la proteína más abundante con 20 millones de copias por célula. En el otro extremo, el factor de transcripción y oncogén FOS está presente en aproximadamente 4.000 copias en la muestra de células HeLa de los inventores. Como se esperaba, las proteínas implicadas en la señalización celular se expresan en general a valores más bajos, como ejemplo incluso el factor de armazón

15 proteína de armazón de proteína quinasa activado por mitógeno 1 (MAP2K1IP1) está presente en solamente 160.000 copias. Sin embargo, se expresan en muy alta medida factores de señalización ubicuos con un papel de tipo chaperona, tales como isoformas 14-3-3 (14-3-3 sigma; 2,1 millones de copias). Dos miembros del ribosoma mitocondrial tienen aproximadamente 200.000 copias en esta línea celular (L23 y L5), mientras que un tercero (L35) tiene aproximadamente 500.000 (obsérvese que no todas las subunidades proteicas ribosómicas tienen estequiometría igual). El genoma mitocondrial solamente codifica 13 genes por tanto es quizás sorprendente que las

20 proteínas implicadas en su traducción sean necesarias en números de copias tan altos. Un miembro de la cadena respiratoria, ATP5B, tiene aproximadamente 5 millones de copias por células HeLa, aproximadamente cinco veces

mayor que PSMC3, un componente regulador del proteasoma. El complejo T es un miembro de un sistema de chaperona y como se esperaba tiene un número de copias muy alto (aproximadamente 4 millones). La ácido graso sintasa, una enzima clásica, se expresa a 3,4 millones de copias, mientras que otra enzima acil coenzima A tioéster hidrolasa (ACOT7) se expresa aproximadamente siete veces menos (450.000 copias). Dichos números de expresión podrían ser interesantes para modelización de rutas metabólicas. Estos son ejemplos anecdóticos pero ilustran que el conocimiento de los niveles de expresión absolutos de proteínas celulares puede contribuir al entendimiento de sus papeles en la célula.

**Tabla Complementaria 3: enlace de UniProt correspondiente para las 43 proteínas usadas para la determinación del número de copias de proteínas.**

Nombres de proteínas	Nombres de genes	de Función
Proteína sigma 14-3-3	SFN	<a href="http://www.uniprot.org/uniprot/P31947">http://www.uniprot.org/uniprot/P31947</a>
Subunidad 6A reguladora de proteasa 26S	PSMC3	<a href="http://www.uniprot.org/uniprot/P17980">http://www.uniprot.org/uniprot/P17980</a>
Proteína ribosómica 28S S23, mitocondrial	MRPS23	<a href="http://www.uniprot.org/uniprot/Q9Y3D9">http://www.uniprot.org/uniprot/Q9Y3D9</a>
Proteína ribosómica 28S S35, mitocondrial	MRPS28	<a href="http://www.uniprot.org/uniprot/Q9Y2Q9">http://www.uniprot.org/uniprot/Q9Y2Q9</a>
Proteína ribosómica 35S L50, mitocondrial	MRPL50	<a href="http://www.uniprot.org/uniprot/Q8N5N7">http://www.uniprot.org/uniprot/Q8N5N7</a>
Proteína de tipo AFG3 2	AFG3L2	<a href="http://www.uniprot.org/uniprot/Q9Y4W6">http://www.uniprot.org/uniprot/Q9Y4W6</a>
Subunidad beta de ATP sintasa, mitocondrial	ATP5B	<a href="http://www.uniprot.org/uniprot/P06576">http://www.uniprot.org/uniprot/P06576</a>
Carbonil reductasa [NADPH] 3	CBR3	<a href="http://www.uniprot.org/uniprot/Q75828">http://www.uniprot.org/uniprot/Q75828</a>
Proteína de cuerpo multivesicular con carga 6	CHMP6	<a href="http://www.uniprot.org/uniprot/Q96FZ7">http://www.uniprot.org/uniprot/Q96FZ7</a>
Subunidad 5 de complejo de signalosoma COP9	COPS5	<a href="http://www.uniprot.org/uniprot/Q92905">http://www.uniprot.org/uniprot/Q92905</a>
Citocromo b5 reductasa 4	CYBSR4	<a href="http://www.uniprot.org/uniprot/Q7L1T6">http://www.uniprot.org/uniprot/Q7L1T6</a>
Subunidad 1 de complejo de citocromo b-c1, mitocondrial	UQCRC1	<a href="http://www.uniprot.org/uniprot/P31930">http://www.uniprot.org/uniprot/P31930</a>
Acil coenzima A tioéster hidrolasa citosólica	ACOT7	<a href="http://www.uniprot.org/uniprot/O00154">http://www.uniprot.org/uniprot/O00154</a>
Proteína asociada a balsa lipídica de retículo endoplásmico 2	ERLIN2	<a href="http://www.uniprot.org/uniprot/O94905">http://www.uniprot.org/uniprot/O94905</a>
Enoil-CoA hidratasa, mitocondrial	ECHS1	<a href="http://www.uniprot.org/uniprot/P30084">http://www.uniprot.org/uniprot/P30084</a>
Factor de inicio de la traducción eucariota 3 subunidad 6	EIF3E	<a href="http://www.uniprot.org/uniprot/P60228">http://www.uniprot.org/uniprot/P60228</a>
Subunidad SSPR1 del complejo de FACT	SSRP1	<a href="http://www.uniprot.org/uniprot/Q08945">http://www.uniprot.org/uniprot/Q08945</a>
Ácido graso sintasa	FASN	<a href="http://www.uniprot.org/uniprot/Q6PJJ3">http://www.uniprot.org/uniprot/Q6PJJ3</a>
Endonucleasa Flap1	FEN1	<a href="http://www.uniprot.org/uniprot/P39748">http://www.uniprot.org/uniprot/P39748</a>
Proteína de choque térmico de 70 kDa 4	HSPA4	<a href="http://www.uniprot.org/uniprot/P34932">http://www.uniprot.org/uniprot/P34932</a>
Antígeno asociado a carcinoma hepatocelular 59	C9orf78	<a href="http://www.uniprot.org/uniprot/Q9NZ63">http://www.uniprot.org/uniprot/Q9NZ63</a>
Proteína de armazón de proteína quinasa activada por mitógeno 1	MAP2K1IP1	<a href="http://www.uniprot.org/uniprot/Q9UHA4">http://www.uniprot.org/uniprot/Q9UHA4</a>
Nucleoproteína TPR	TPR	<a href="http://www.uniprot.org/uniprot/P12270">http://www.uniprot.org/uniprot/P12270</a>
Peptidil-propil cis-trans isomerasa B	PPIB	<a href="http://www.uniprot.org/uniprot/P23284">http://www.uniprot.org/uniprot/P23284</a>
Peroxisredoxina 6	PRDX6	<a href="http://www.uniprot.org/uniprot/P30041">http://www.uniprot.org/uniprot/P30041</a>
Subunidad de prefoldina 1	PFDN1	<a href="http://www.uniprot.org/uniprot/O60925">http://www.uniprot.org/uniprot/O60925</a>
Regulador de corte y empalme pre ARNm WTAP	WTAP	<a href="http://www.uniprot.org/uniprot/Q15007">http://www.uniprot.org/uniprot/Q15007</a>
ARN helicasa dependiente de ATP probable DDX20	DDX20	<a href="http://www.uniprot.org/uniprot/Q9UHI6">http://www.uniprot.org/uniprot/Q9UHI6</a>
Protooncogén C-Fos	FOS	<a href="http://www.uniprot.org/uniprot/P01100">http://www.uniprot.org/uniprot/P01100</a>
Nucleósido fosforilasa de purina	NP	<a href="http://www.uniprot.org/uniprot/P00491">http://www.uniprot.org/uniprot/P00491</a>
Proteína de tipo activadora de GTPasa Ras IQGAP1	IQGAP1	<a href="http://www.uniprot.org/uniprot/P46940">http://www.uniprot.org/uniprot/P46940</a>
Proteína de unión a ARN de interacción tallo-bucle SRA, mitocondrial	C14orf156	<a href="http://www.uniprot.org/uniprot/Q9GZT3">http://www.uniprot.org/uniprot/Q9GZT3</a>
Subunidad beta de proteína de complejo T 1	CCT2	<a href="http://www.uniprot.org/uniprot/P78371">http://www.uniprot.org/uniprot/P78371</a>
Subunidad 1 de complejo THO	THOC1	<a href="http://www.uniprot.org/uniprot/Q96FV9">http://www.uniprot.org/uniprot/Q96FV9</a>
ATPasa de retículo endoplásmico transicional	VCP	<a href="http://www.uniprot.org/uniprot/P55072">http://www.uniprot.org/uniprot/P55072</a>
Vimentina	VIM	<a href="http://www.uniprot.org/uniprot/P08670">http://www.uniprot.org/uniprot/P08670</a>
Proteína de dedo de zinc 828	C13orf8	<a href="http://www.uniprot.org/uniprot/Q96JM3">http://www.uniprot.org/uniprot/Q96JM3</a>

*Cuantificación absoluta usando PrEST pesados* – anteriormente se han usado PrEST ya expresados y purificados y se han cuantificado frente a proteína ABP pesada y lisado celular marcado con SILAC pesado. Aunque es conveniente determinar los números de copias en líneas celulares, en otras aplicaciones sería más apropiado expresar PrEST marcados pesados, que pueden después mezclarse en cualquier proteoma elegido, incluyendo muestras tisulares y de fluido corporal clínico. Para aplicar el enfoque de cuantificación absoluta a muestras no marcadas se expresaron 28 de los PrEST en *E. coli* marcado con SILAC pesado, se purificaron y se preparó una mezcla maestra pesada. Para simplificar la cuantificación de niveles de PrEST, se desarrolló un conjunto automático que empleaba nanoelectropulverización estática (Advion NanoMate; véase Ejemplo 1). Como se esperaba, la adición de la mezcla maestra pesada a células normales, no marcadas con SILAC permitió una cuantificación igualmente sencilla de las proteínas diana con buena correlación con el experimento previo (Figura 7). Se proporciona información detallada acerca de la identificación y cuantificación de las proteínas en la Tabla Complementaria 4.

**Tabla Complementaria 4a:** toda la información de identificación y cuantificación del experimento en el que se añadieron PrEST pesados a lisado de HeLa no marcado

Nombres de proteínas	Nombre gen	Secuencia del	Relación H/L de la mezcla			Conc. De proteína µg/µl	PrEST (1+2)
			1	2	3		
Proteína de tipo AFG3 2	AFG3L2	EQLYTK	NaN	NaN	NaN	0,779	0,01284
Proteína de tipo AFG3 2	AFG3L2	HFEQAJER	NaN	NaN	1,0384	0,779	0,01284
Proteína de tipo AFG3 2	AFG3L2	HLSDSINQK	0,98061	NaN	1,034NaN	0,779	0,01284
Proteína de tipo AFG3 2	AFG3L2	LASLTPGFSGADVAVNCN	0,64771	0,68664	0,73842	0,779	0,01284
Proteína de tipo AFG3 2	AFG3L2	TVAYHEAGHAVAGWYLEH	0,99036	0,88405	1,0667	0,779	0,01284
		ADPLLK					
Proteína de tipo AFG3 2	AFG3L2	VSEEIFFR	0,88034	1,0026	0,087458	0,779	0,01284
Subunidad beta de ATP sintasa	ATP5B	IMNVIGEPIDER	0,49547	0,38751	0,46817	0,359	0,25036
Subunidad beta de ATP sintasa	ATP5B	IPVGPETLGR	0,67213	0,52947	0,5949	0,359	0,25036
Subunidad beta de ATP sintasa	ATP5B	LVEVAQHLGESTVR	0,57216	0,53836	0,64057	0,359	0,25036
Subunidad beta de ATP sintasa	ATP5B	TIAMDGTEGLVR	0,52173	0,3056	0,41416	0,359	0,25036
Subunidad beta de ATP sintasa	ATP5B	VLDSGAPIK	0,69347	0,569	0,60973	0,359	0,25036
Subunidad beta de ATP sintasa	ATP5B	VLDSGAPIKIPVGPETLGR	0,71587	0,67498NaN		0,359	0,25036
Carbonil reductasa [NADPH]	CBR3	AFENCSEDLQER	3,7066	3,6554	2,7849	0,692	0,01445
Carbonil reductasa [NADPH]	CBR3	FHSETLTEGDLVDLMK	0,82772	0,54401	0,65383	0,692	0,01445
Carbonil reductasa [NADPH]	CBR3	WNISLQCLR	NaN	0,61063	0,027081	0,692	0,01445
Subunidad de proteína de complejo T 1	de CCT2	HGINCFINR	0,96794	0,9253	0,80254	0,392	0,38219
Subunidad de proteína de complejo T 1	de CCT2	LIANTGMDTDK	0,86038	0,52336	0,58541	0,392	0,38219
Subunidad de proteína de complejo T 1	de CCT2	LALVTGGEIASTFDHPPELV	1,0899	0,99631	1,1095	0,392	0,38219
Subunidad de proteína de complejo T 1	de CCT2	LIEEVMIGEDK	0,83525	0,94634	0,78472	0,392	0,38219
Subunidad de proteína de complejo T 1	de CCT2	VAEIEHAEK	1,1668	0,99657	1,1107	0,392	0,38219
Subunidad de proteína de complejo T 1	de CCT2	VAEIEHAEKEK	0,84991	0,99639NaN		0,392	0,38219
Complejo de signalosoma COP9	COP55	DHMYFK	0,72801	0,3227	0,37943	0,999	0,01001
Complejo de signalosoma COP9	COP55	ISALALLK	NaN	NaN	0,23709	0,999	0,01001
Citocromo b5 reductasa 4	CYB5R4	QGHISPELLSEFLK	NaN	12,575	13,552	0,867	0,01153
Citocromo b5 reductasa 4	CYB5R4	TEDDIWR	11,628	12,424	12,299	0,867	0,01153
R dependiente de AT probable	DDX20	VLISDITSR	0,59562NaN	NaN		0,705	0,01419
Enoil-CoA hidratasa mitocondrial	ECHS1	EGMTAFVEK	1,0475NaN		0,028231	0,342	0,58542
Enoil-CoA hidratasa mitocondrial	ECHS1	ESVNAAFEMTLTEGSK	1,0239NaN		0,98226	0,342	0,58542
Enoil-CoA hidratasa mitocondrial	ECHS1	ICPVETLVEEAIQCAEK	0,77883	0,73163	0,73992	0,342	0,58542

**Tabla Complementaria 4a (continuación):**

µl PrESTs	NanoMate pmol/µl	NanoMate pmol/µl	NanoMate pmol/µl	pmol PrEST 1	pmol PrEST 2	pmol PrEST 3	pmol Proteína 1	pmol Proteína 2	pmol Proteína 3
0,01284	18,253386	22,899283		0,2343735	0,294027	0	0,239008	0,284359	0
0,01284	18,253386	22,899283		0,2343735	0,294027	0	0,361849	0,428211	0
0,01284	18,253386	22,899283		0,2343735	0,294027	0	0,236655	0,332591	0
0,01284	18,253386	22,899283		0,2343735	0,294027	0	0,266231	0,293264	0
0,01284	18,253386	22,899283	19,970078	0,2343735	0,294027	0	12,09394	14,98731	10,67926
0,25036	23,934268	23,197529	19,970078	5,9921835	5,807733	4,999709	8,915215	10,96896	8,404284
0,25036	23,934268	23,197529	19,970078	5,9921835	5,807733	4,999709	10,47292	10,78782	7,805093
0,25036	23,934268	23,197529	19,970078	5,9921835	5,807733	4,999709	11,48522	19,00436	12,07193
0,25036	23,934268	23,197529	19,970078	5,9921835	5,807733	4,999709	10,640869	10,20691	8,199873
0,01445	71,889794	33,419567	19,970078	5,9921835	5,807733	4,999709	8,370491	8,604304	
0,01445	71,889794	33,419567	44,296682	1,0388075	0,482913	0,640087	0,280259	0,132109	0,229842
0,01445	71,889794	33,419567	44,296682	1,0388075	0,482913	0,640087	1,255023	0,887691	0,978981
0,01445	71,889794	33,419567	44,296682	1,0388075	0,482913	0,640087	0,790843	0,790843	23,63602
0,38219	17,45966	17,45966	12,336037	0	6,661102	4,71471	0	7,198856	5,874735
0,38219	17,45966	17,45966	12,336037	0	6,672907	4,71471	0	12,75013	8,053689
0,38219	17,45966	17,45966	12,336037	0	6,672907	4,71471	0	6,697622	4,2494
0,38219	17,45966	17,45966	12,336037	0	6,672907	4,71471	0	7,051279	6,008143
0,38219	17,45966	17,45966	12,336037	0	6,672907	4,71471	0	6,695874	4,244809
0,01001	14,240375	16,354906	12,336037	0,1425462	0,163713	0	0,195802	0,507321	0
0,01001	14,240375	16,354906	16,354906	0,1425462	0,163713	0	0	0	0
0,01153	21,356568	20,719244	20,719244	0,2462412	0,238893	0	0,021177	0,018997	0
0,01153	21,356568	20,719244	20,719244	0,2462412	0,238893	0	0,021177	0,019228	0
0,01419	20,843513	13,238853	14,941964	0,2957695	0,187859	0,212026	0,496574		156,5016
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	4,033307		4,49799
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	4,126271		5,97118
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	5,424661	10,9591	

**Tabla Complementaria 4a (continuación):**

Nombres de proteínas	Nombre de gen	Secuencia	Relación mezcla 1	H/L deRelación maestramezcla 2	H/L deRelación maestramezcla 3	Relación H/L de maestramezcla maestra	Conc. De proteína µg/µl	PrEST (1+2)
Enoil-CoA hidratasa, mitocondrial	ECHS1	IVVAMAK	NaN	NaN	1,5022	1,063	0,342	0,58542
Enoil-CoA hidratasa, mitocondrial	ECHS1	LFYSTFATDDR	NaN	1,5764	1,5022	1,6594	0,342	0,58542
Enoil-CoA hidratasa, mitocondrial	ECHS1	LFYSTFATDDRK	NaN	NaN	1,391NaN	0,342	0,58542	0,58542
Enoil-CoA hidratasa, mitocondrial	ECHS1	SLAMEMVLGTGR	NaN	0,5035	0,63232	0,092059	0,342	0,58542
Inicio de la traducción eucariota	EIF3E	LGHWMGNNAVSPYQQVIEK	NaN	0,11968	0,12455	0,048128	0,714	0,01401
Inicio de la traducción eucariota	EIF3E	LNMTPEEAER	NaN	0,096546	0,088999NaN	0,081968	0,714	0,01401
Inicio de la traducción eucariota	EIF3E	SQMLAMNIEK	NaN	0,1375	0,096998	0,082782	0,714	0,01401
Inicio de la traducción eucariota	EIF3E	WIVNLIJR	NaN	0,50502	2,1428	0,19075	0,186	0,05364
Lípido de retículo endoplásmico	ERLIN2	ADAECYTAMK	NaN	4,6716	0,36952	0,33782	0,186	0,05364
Lípido de retículo endoplásmico	ERLIN2	DIPNMFMSAGSVSK	NaN	0,76436	5,1738	4,6718	0,186	0,05364
Lípido de retículo endoplásmico	ERLIN2	LSFGLEDEPLETATK	NaN	0,55946	0,99457	0,95839	0,186	0,05364
Lípido de retículo endoplásmico	ERLIN2	LTPEYLQLMK	NaN	0,57294NaN	3,8443	4,775	0,186	0,05364
Endonucleasa Flap 1	FEN1	VAQVAEITYGQK	NaN	0,53151	0,53151	0,62173	0,883	0,07927
Endonucleasa Flap 1	FEN1	EAHQFLEPEVLDPESVELK	NaN	NaN	NaN	0,64887	0,883	0,26011
Endonucleasa Flap 1	FEN1	HLTASEAK	NaN	0,45538	0,50335	0,51855	0,883	0,07927
Endonucleasa Flap 1	FEN1	KLPIQEFHLSR	NaN	0,53912	0,53912	0,63616	0,883	0,07927
Endonucleasa Flap 1	FEN1	LDPNKYPVPENWLHK	NaN	0,56689	0,56689	0,65691	0,883	0,07927
Endonucleasa Flap 1	FEN1	LPIQEFHLSR	NaN	0,46634	0,2953 INaN	0,2507	0,883	0,07927
Endonucleasa Flap 1	FEN1	SIEEIVR	NaN	0,25573	0,23882	0,2507	0,883	0,07927
Endonucleasa Flap 1	FEN1	VYAAATEDMDCLTFGSPVLMR	NaN	0,48958	0,15763	0,65879	0,883	0,07927
Endonucleasa Flap 1	FEN1	YVPENWLHK	NaN	NaN	NaN	0,64887	0,769	0,26011
Proteína de choque térmico de 70 kDa	HSPA4	EDQYDHLDAADMTK	NaN	NaN	NaN	0,64887	0,769	0,26011
Proteína de choque térmico de 70 kDa	HSPA4	LNLQNK	NaN	2,3137NaN	2,1319	2,1319	0,769	0,26011
Proteína de choque térmico de 70 kDa	HSPA4	NKEDQYDHLDAADMTK	NaN	0,76949	0,53896	0,54666	0,769	0,26011
Proteína de choque térmico de 70 kDa	HSPA4	QSLTMDPWK	NaN	0,50413	0,4421	0,54603	0,769	0,26011
Proteína de choque térmico de 70 kDa	HSPA4	STNEAMEWMNKK	NaN	0,056431	0,045314	0,041474	0,769	0,26011
Proteína ribosómica 39S L50	MRPL50	AYTPPEDLQSR	NaN	0,65073NaN	1NaN	0,684	0,684	0,01462
Proteína ribosómica 39S L50	MRPL50	LESYVK	NaN	1,131NaN	1NaN	0,684	0,684	0,01462
Proteína ribosómica 28S S23	MRPS23	ALLAEGVILR	NaN	0,91056	0,91056	0,4868	0,407	0,02459
Proteína ribosómica 28S S23	MRPS23	LFVETGK	NaN	1,0838NaN	1,0838NaN	0,407	0,407	0,02459
Proteína ribosómica 28S S23	MRPS23	LGETDEEK	NaN	0,93474	1,0182	1,0733	0,407	0,02459

Tabla Complementaria 4a (continuación):

µl PrESTs	NanoMate pmol/µl	NanoMate pmol/µl	NanoMate pmol/µl	pmol PrEST 1	pmol PrEST 2	pmol PrEST 3	pmol Proteína 1	pmol Proteína 2	pmol Proteína 3
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	2,680087	5,337511	4,156346
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	2,680087	5,337511	2,662526
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	8,391041	5,764204	47,99309
0,58542	7,2168512	13,696164	7,5470525	4,224889	8,018008	4,418195	2,267498	12,6803	5,103105
0,01401	19,370034	20,330459	17,530494	0,2713742	0,28483	0,245602	3,20037	2,286871	3,20037
0,01401	19,370034	20,330459	17,530494	0,2713742	0,28483	0,245602	2,810828	2,936449	2,996318
0,01401	19,370034	20,330459	17,530494	0,2713742	0,28483	0,245602	2,071489	2,071489	2,966855
0,01401	19,370034	20,330459	17,530494	0,2713742	0,28483	0,245602	0,613804	0,613804	0,621326
0,05364	24,33193	24,52011	22,095083	1,3051647	1,315259	1,18518	2,584382	3,559371	3,508319
0,05364	24,33193	24,52011	22,095083	1,3051647	1,315259	1,18518	0,279383	0,254215	0,253688
0,05364	24,33193	24,52011	22,095083	1,3051647	1,315259	1,18518	1,707526	1,32244	1,236637
0,05364	24,33193	24,52011	22,095083	1,3051647	1,315259	1,18518	0,342132	0,342132	0,248205
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	2,910183	3,067057	2,090487
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	2,841713	3,238644	2,506447
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	3,575323	3,023763	2,043068
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	2,87564	2,87564	1,978533
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	3,491296	5,52039	5,184357
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	6,366601	6,825941	1,972887
0,07927	20,539053	20,564795	16,396092	1,6281308	1,630171	1,299718	3,325566	10,34176	9,973034
0,26011	32,693479	32,693479	24,878715	0	8,503901	6,471203	0	15,77835	3,035416
0,26011	32,693479	32,693479	24,878715	0	8,503901	6,471203	0	19,23524	11,83771
0,26011	32,693479	32,693479	24,878715	0	8,503901	6,471203	0	187,6661	11,85137
0,26011	32,693479	32,693479	24,878715	0	8,503901	6,471203	0	0	156,0303
0,01402	23,657325	15,986181	18,15397	0,3458701	0,233718	0,254519	0,531511	0,416141	0,650264
0,01402	23,657325	15,986181	18,15397	0,3458701	0,233718	0,254519	0,305809	0,349623	0,349623
0,02459	14,832962	15,409572	12,873057	0,3647425	0,378921	0,316548	0,316548	0,372148	0,29493
0,02459	14,832962	15,409572	12,873057	0,3647425	0,378921	0,316548	0,390207	0,372148	0,29493
0,02459	14,832962	15,409572	12,873057	0,3647425	0,378921	0,316548	0,390207	0,372148	0,29493

Tabla Complementaria 4a (continuación):

Nombres de proteínas	Nombre de gen	Secuencia	Relación mezcla 1	H/L de mezcla 1	Relación H/L de mezcla 2	H/L de mezcla 2	Relación H/L de mezcla 3	H/L de mezcla 3	Conc. De proteína µg/µl	PrEST (1+2)
Proteína ribosómica 28S S23,	MRPS23	TQHGGSHVSR	NaN	1,0302	NaN	1,1676	NaN	0,407	0,02459	
Proteína ribosómica 28S S23,	MRPS23	YTELQK	NaN	NaN	NaN	NaN	NaN	0,407	0,02459	
Proteína ribosómica 28S S28,	MRPS28	AGGFASALER	NaN	0,51434	NaN	NaN	NaN	0,449	0,02229	
Proteína ribosómica 28S S28,	MRPS28	HSELLQK	NaN	NaN	NaN	NaN	NaN	0,449	0,02229	
Nucleósido purina fosfor	NP	ACVMMQGR	0,58177	0,63931	0,56821	0,56821	0,56821	1,065	0,09389	
Nucleósido purina fosfor	NP	DHINLPFGSQNPLR	0,50497	0,985	0,93693	0,93693	0,93693	1,065	0,09389	
Nucleósido purina fosfor	NP	FEVGD1MLIR	0,68963	0,70356	0,71547	0,71547	0,71547	1,065	0,09389	
Nucleósido purina fosfor	NP	FHMYEGYPLWK	0,72381	0,90091	0,76466	0,76466	0,76466	1,065	0,09389	
Nucleósido purina fosfor	NP	HRPQVAICGSLGGLTDK	NaN	0,85363	0,81538	0,81538	0,81538	1,065	0,09389	
Nucleósido purina fosfor	NP	LTQAQIFDYGEIPNFR	1,118	1,8464	1,8793	1,8793	1,8793	1,065	0,09389	
Nucleósido purina fosfor	NP	LVFGLNGR	1,1247	0,30888	1,2496	1,2496	1,2496	1,065	0,09389	
Nucleósido purina fosfor	NP	STVPGHAGR	1,0994	NaN	0,49901	0,49901	0,49901	1,065	0,09389	
Nucleósido purina fosfor	NP	VFHLLGVDTLWTNAAGLNPK	NaN	1,0676	1,0636	1,0636	1,0636	1,065	0,09389	
Poli [ADP-ribosa] polimerasa	PARP4	AEGILLLVK	NaN	0,58019	0,44594	0,44594	0,44594	0,579	0,01727	
Subunidad de prefolidina 1	PFDN1	EAIHSQLEK	0,76453	NaN	NaN	NaN	NaN	0,441	0,09078	
Subunidad de prefolidina 1	PFDN1	LADIQEQLNR	1,6259	1,8677	1,9142	1,9142	1,9142	0,441	0,09078	
Subunidad de prefolidina 1	PFDN1	MFILQSK	1,1344	NaN	NaN	NaN	NaN	0,441	0,09078	
Peroxirredoxina 6	PRDX6	DFTPVCTTELGR	0,56628	NaN	0,33394	0,33394	0,33394	0,714	0,56054	
Peroxirredoxina 6	PRDX6	DINAYNCEEPTEK	0,61696	0,60109	0,5663	0,5663	0,5663	0,714	0,56054	
Peroxirredoxina 6	PRDX6	ELAILGMLDPAEK	0,57928	0,38932	0,38813	0,38813	0,38813	0,714	0,56054	
Peroxirredoxina 6	PRDX6	ELAILGMLDPAEKDEK	0,56863	0,74994	0,21052	0,21052	0,21052	0,714	0,56054	
Peroxirredoxina 6	PRDX6	FHDFLGDSWGLFSPR	NaN	0,57213	0,59897	0,59897	0,59897	0,714	0,56054	
Peroxirredoxina 6	PRDX6	GMPVTAR	NaN	0,59916	0,51841	0,51841	0,51841	0,714	0,56054	
Peroxirredoxina 6	PRDX6	LAPEFAK	0,80521	0,6713	0,74114	0,74114	0,74114	0,714	0,56054	
Peroxirredoxina 6	PRDX6	LIALSIDSVEDHLAWSK	0,74873	0,66939	0,62365	0,62365	0,62365	0,714	0,56054	
Peroxirredoxina 6	PRDX6	LPFPIDDR	0,66166	0,68947	0,66519	0,66519	0,66519	0,714	0,56054	
Peroxirredoxina 6	PRDX6	WVFGPDK	0,52861	0,25261	0,63133	0,63133	0,63133	0,714	0,56054	
Peroxirredoxina 6	PRDX6	WVFGPDKK	0,84361	0,36625	0,63578	0,63578	0,63578	0,714	0,56054	
Subunidad reguladora de proteasa 26S	PSMC3	AVCVEAGMIALR	NaN	0,27688	0,06781	0,06781	0,06781	0,672	0,04466	
Subunidad reguladora de proteasa 26S	PSMC3	GATELTHEDYMEGILEVQAK	NaN	NaN	NaN	NaN	NaN	0,672	0,04466	
Subunidad reguladora de proteasa 26S	PSMC3	MNVSPDVNYEELAR	0,67529	0,53568	0,63201	0,63201	0,63201	0,672	0,04466	
Subunidad SSRP de complejo FACT	SSRP1	ADVIGATGDAICIFR	0,51416	0,56819	0,54776	0,54776	0,54776	0,587	0,05111	

**Tabla Complementaria 4a (continuación):**

µl PrESTs	NanoMate pmol/µl	NanoMate pmol/µl	NanoMate pmol/µl	pmol PrEST 1	pmol PrEST 2	pmol PrEST 3	pmol Proteína 1	pmol Proteína 2	pmol Proteína 3
0,02459	14,832962	15,409572	12,873057	0,3647425	0,378921	0,316548		0,367813	
0,02459	14,832962	15,409572	12,873057	0,3647425	0,378921	0,316548			0,27111
0,02229	11,385909	11,881033	10,583484	0,2537919	0,264828	0,235906		0,514889	
0,02229	11,385909	11,881033	10,583484	0,2537919	0,264828	0,235906	0,436241		
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	3,361868	4,03268	3,445905
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	1,779425	2,617394	2,089801
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	2,461671	3,664411	2,736659
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	2,345425	2,861699	2,560612
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	1,518464	3,020199	2,401331
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	1,509418	1,396302	1,041876
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998	1,544154	8,346714	1,566899
0,09389	18,081184	27,459079	20,854165	1,6976424	2,578133	1,957998		2,414887	3,923764
0,01727	19,772134	22,931238	20,854165	1,6976424	2,578133	1,957998		0,682574	1,840915
0,09078	16,010502	16,223891	1,4534334	0,3414648	0,396022	0	1,901081		0
0,09078	16,010502	16,223891	1,4534334	1,472805	1,472805	0	0,893925	0,788566	
0,09078	16,010502	16,223891	1,4534334	1,472805	1,472805	0	1,281235		
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	19,14733	18,5608	21,14787
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	17,57448	18,5608	12,47063
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	18,71764	28,65692	18,19524
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	19,0682	14,8768	33,54607
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119		19,50031	11,79044
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119		18,62059	13,62265
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	13,46574	16,61956	9,528724
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	14,48153	16,66698	11,32385
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	16,3872	16,18158	10,61669
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	20,51182	44,16575	11,1861
0,56054	19,343405	19,903505	12,598777	10,842752	11,15671	7,062119	12,8528	30,46201	11,1078
0,04466	34,506036	32,565659	25,92524	1,5410396	1,454382	1,157821		5,252753	17,07374
0,04466	34,506036	32,565659	25,92524	1,5410396	1,454382	1,157821	2,282041	2,715021	1,831967
0,04466	34,506036	32,565659	25,92524	1,5410396	1,454382	1,157821	1,856034	1,525648	1,240508
0,05111	18,671463	16,96063	13,294863	0,9542985	0,866858	0,6795			

Tabla Complementaria 4a (continuación):

Nombres de proteínas	Nombre de gen	Secuencia	Relación H/L mezcla maestra 1	Relación H/L mezcla maestra 2	Relación H/L mezcla maestra 3	Conc. de proteína µg/µl	PrEST (1+2)
Subunidad SSRP de complejo FACT	SSRP1	ELQLTPR	NaN	0,48128	NaN	0,587	0,05111
Subunidad SSRP de complejo FACT	SSRP1	IPYTTVLR	NaN	0,57615	NaN	0,587	0,05111
Subunidad 1 de complejo THO	THOC1	AVNNSNYGWR	NaN	NaN	0,96021	0,953	0,0105
Subunidad 1 de complejo THO	THOC1	LWNLCPDNMEACK	NaN	0,064595	0,05542	0,953	0,0105
Subunidad 1 de complejo THO	THOC1	SLPEYLENMVVK	NaN	0,063485	0,08477	0,953	0,0105
Subunidad 1 de complejo THO	THOC1	TGEDEEEDNDALUC	NaN	2,2196	1,3318	0,953	0,0105
Nucleoproteína TPR	TPR	ILLSQTTGVAIPLHASSLDDVSLSTPK	NaN	2,6647	2,8484	0,388	0,07732
Nucleoproteína TPR	TPR	ITEIQLK	NaN	NaN	2,4635	0,388	0,07732
Nucleoproteína TPR	TPR	LESALTELEQLR	NaN	2,4601	2,3322	0,388	0,07732
Nucleoproteína TPR	TPR	LESALTELEQLRK	NaN	NaN	2,7043	0,388	0,07732
Nucleoproteína TPR	TPR	NIEELQQNQNR	NaN	2,2529	2,0753	0,388	0,07732
Nucleoproteína TPR	TPR	QHQMQLVDSIVR	NaN	NaN	1,9585	0,388	0,07732
Complejo s de citocromo b-c1	UQCRC1	ADLTEYLSHYK	NaN	0,21768	0,25733	0,43	0,02326
Complejo s de citocromo b-c1	UQCRC1	DVVFNYLHATAFQGTPLAQAVEGPFSENVR	NaN	0,20672	0,25427	0,43	0,02326
Complejo s de citocromo b-c1	UQCRC1	MVLAAGGVEHQQLDLAQK	NaN	0,036719	0,16497	0,43	0,02326
Vimentina	VIM	DNLAEIDIMR	NaN	0,38208	0,423	0,427	1,17023
Vimentina	VIM	EEAENTLQSFRR	NaN	0,53572	0,51411	0,427	1,17023
Vimentina	VIM	EKLQEEMLQR	NaN	0,35844	0,35626	0,427	1,17023
Vimentina	VIM	ILLAELEQLK	NaN	0,60792	0,60021	0,427	1,17023
Vimentina	VIM	ILLAELEQLKGQGGK	NaN	0,45207	NaN	0,427	1,17023
Vimentina	VIM	LGDLYEEEMR	NaN	0,41358	0,39183	0,427	1,17023
Vimentina	VIM	LQEEMLQR	NaN	0,43217	0,27333	0,427	1,17023
Vimentina	VIM	QDVVDNASLAR	NaN	0,50789	0,50041	0,427	1,17023
Vimentina	VIM	QVDQLTNDK	NaN	0,57376	0,52999	0,427	1,17023
Vimentina	VIM	RQVDQLTNDK	NaN	0,5666	0,51939	0,427	1,17023

**Tabla Complementaria 4a (continuación):**

ul PrESTs (3)	NanoMate pmol/ul	NanoMate pmol/ul	NanoMate pmol/ul	pmol PrEST 1	pmol PrEST 2	pmol PrEST3	pmol Proteina 1	pmol Proteina 2	pmol Proteina 3
0,05111	18,671463	16,96063	13,294863	0,9542985	0,866858	0,6795	1,553296	1,801151	
0,05111	18,671463	16,96063	13,294863	0,9542985	0,866858	0,6795		1,50457	
0,0105	26,314913	19,802314	20,98021	0,2763066	0,207924	0,220292	1,537428	3,218892	0,229421
0,0105	26,314913	19,802314	20,98021	0,2763066	0,207924	0,220292		3,275172	3,974959
0,0105	26,314913	19,802314	20,98021	0,2763066	0,207924	0,220292	0,14464	0,093676	2,598705
0,0105	26,314913	19,802314	20,98021	0,2763066	0,207924	0,220292	0,494341	0,478378	0,165409
0,07732	16,099976	16,486454	15,642843	1,2448502	1,274733	1,209505		0,424626	0,49097
0,07732	16,099976	16,486454	15,642843	1,2448502	1,274733	1,209505	0,528442	0,518163	0,518611
0,07732	16,099976	16,486454	15,642843	1,2448502	1,274733	1,209505		0,447252	0,447252
0,07732	16,099976	16,486454	15,642843	1,2448502	1,274733	1,209505		0,58281	0,58281
0,07732	16,099976	16,486454	15,642843	1,2448502	1,274733	1,209505	0,687307	0,617567	0,617567
0,02326	20,444076	20,444076	15,780782	0	0,475529	0,367061	0	2,184533	1,426421
0,02326	20,444076	20,444076	15,780782	0	0,475529	0,367061	0	2,126601	1,443587
0,02326	20,444076	20,444076	15,780782	0	0,475529	0,367061	0	2,882519	2,363256
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	43,04868	50,0032	45,65171
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	30,70288	41,14169	32,78401
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	45,88785	59,37055	51,04452
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	27,05626	35,23992	29,32827
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	36,38384		
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	39,76991	53,98094	59,59858
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	38,05919	77,38394	47,80201
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	32,38504	42,26805	35,99835
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	28,66711	39,90897	36,10945
1,17023	14,055391	18,074527	15,696811	16,44804	21,15135	18,36888	29,02937	40,72345	27,26971

**Tabla Complementaria 4b:**

Nombre de gen	Nombre de proteína	Mezcla maestra 1 (pmol)	Mezcla maestra 2 (pmol)	Mezcla maestra 3 (pmol)	Número de copias 1	Número de copias 2	Número de copias 3	Mediana	DTR (%)
AFG3L2	Proteína de tipo AFG3 2	0,252619235	0,312927491	0	152.131			152.131	
ATP5B	Subunidad beta de ATP sintasa, mitocondrial	9,694065359	10,87839063	8,404284127	5.837.904	6.551.121	5.061.179	5.837.904	12,81
ATYL2	Lisofosfatidilcolina aciltransferasa	-	-	-	-	-	-	-	-
C1orf65	Proteína no caracterizada C1orf65	-	-	-	-	-	-	-	-
CBR3	Carbonil reductasa [NADPH] 3	0,482912745	0,790843465	0,97898086	290.817	476.257	589.556	476.257	33,35
CCT2	Subunidad beta de proteína de complejo T 1	0	6,874450414	5,87473498		4.139.892	3.537.849	3.838.870	11,09
COPS5	Subunidad de complejo de signalosoma COP9	0,195802466	0,507321377	0	117.915	305.516		211.716	62,66
CYBR4	Citocromo b5 reductasa 4	0,021176576	0,019112892	0	12.753	11.510		12.131	7,24
DDX20	ARN helicasa dependiente de ATP probable	0,49657408	-	-	299.044			299.044	
ECHS1	Enoil-CoA hidratasa, mitocondrial	4,126271164	8,361653639	5,234584905	2.484.899	5.035.506	3.152.341	3.152.341	37,18
EIF3E	Factor de inicio de la traducción eucariota	2,539162992	2,611659997	2,996318259	1.529.120	1.572.779	1.804.425	1.572.779	9,05
ERLIN2	Balsa lipídica de retículo endoplásmico -a	1,707526217	0,613803749	0,621326468	1.028.296	369.641	374.172	374.172	64,16
FEN1	Endonucleasa de Flap 1	3,408430951	3,238643686	2,06677383	2.052.605			2.052.605	
HSPA4	Proteína de choque térmico de 70 kDa 4	0	19,23524295	11,83771012		11.583.736	7.128.837	9.356.286	33,67
MLKL	Tipo dominio de quinasa de linaje mixto	-	-	-	-	-	-	-	-
MRPL50	Proteína ribosómica 39S L50, mitocondrial	0,418660011	-	-	252.123			252.123	
MRPS23	Proteína ribosómica 28S S23, mitocondrial	0,390207475	0,369980833	0,294930107	234.988		177.611	206.300	19,67
MRPS28	Proteína ribosómica 28S S35, mitocondrial	0,436240972	0,441351419	0,271110383	262.710	265.788	163.267	262.710	25,29
NP	Purina nucleosido fosforilasa	1,779424739	2,940949047	2,401331339	1.071.595	1.771.081	1.446.116	1.446.116	24,48
PARP4	Poli [ADP-ribosa] polimerasa 4	0,682573768	0	0		411.056		411.056	
PFDN1	Subunidad de prefolidina 1	1,281235376	0,788566037	0	771.578			771.578	
PRDX6	Peroxirredoxina 6	17,57448193	18,59069261	11,79043804	10.583.602	11.195.579	7.100.369	10.583.602	22,95
PSMC3	Subunidad 6 reguladora de proteasa 26S	2,282041145	3,983887093	9,452851115	1.374.278	2.399.153	5.692.641	2.399.153	71,51
SSRP1	Subunidad SSRP1 de complejo FACT	1,704665071	1,5256647761	1,240507611	1.026.573	918.767	747.051	918.767	15,71
THOC1	Subunidad 1 de complejo THO	0,841034292	3,218891514	1,414062794	506.483		851.569	679.026	35,94
TPR	Nucleoproteína TPR	0,528441724	0,518162926	0,504790509	318.235	312.045	303.992	312.045	2,29
UQCRC1	Subunidad de complejo de citocromo b-c1	0	2,184533321	1,443587459		1.315.557	869.349	1.092.453	28,88
VIM	Vimentina	34,384444006	42,26804731	36,10945422	20.706.797	25.454.417	21.745.625	21.745.625	11,03

**Tabla Complementaria 4c:**

Nombre gen	deNombre de proteína	Experimento directo (PrEST ligero)				DTR (%)
		Número de copias 1	Número de copias 2	Número de copias 3	Mediana	
AFG3L2	Proteína de tipo AFG3 2	369.737	412.509	165.983	369.737	41,68
ATP5B	Subunidad beta de ATP sintasa, mitocondrial	5.672.473	4.376.424	4.511.967	4.511.967	14,68
AYTL2	Lisofosfatidilcolina aciltransferasa 1					
C1orf65	Proteína no caracterizada C1orf65	79.823	61.399	322.454	79.823	94,26
CBR3	Carbonil reductasa [NADPH] 3	7.447.762	2.757.533	4.479.130	4.479.130	48,47
CCT2	Subunidad beta de proteína de complejo T 1	323.791	284.218	435.937	323.791	22,62
COPS5	Subunidad de complejo de señalosoma COP9	16.205	10.180	9.515	10.180	30,8
CYB5R4	Citocromo b5 reductasa 4	242.403	184.529		213.466	19,17
DDX20	ARN helicasa dependiente de ATP probable	2.965.394	1.723.133	2.105.336	2.105.336	28,1
ECHS1	Enoil-CoA hidratasa, mitocondrial	1.067.627	599.306	1.253.469	1.067.627	34,63
EIF3E	Factor de inicio de la traducción eucariota	206.262	148.785	149.867	149.867	19,53
ERLIN2	Balsa lipídica de retículo endoplásmico -a	2.372.346	2.019.699	1.563.785	2.019.699	20,42
FEN1	Endonucleasa de Flap 1	2.146.713	1.499.858	1.646.549	1.646.549	19,22
HSPA4	Proteína de choque térmico de 70 kDa 4	128.711		100.891	114.801	17,14
MLKL	Tipo dominio de quinasa de linaje mixto	177.937	250.001	194.935	194.935	18,14
MRPL50	Proteína ribosómica 39S L50, mitocondrial	223.198	203.672	282.020	223.198	17,26
MRPS23	Proteína ribosómica 28S S23, mitocondrial'	473.409	284.783	422.825	422.825	24,8
MRPS28	Proteína ribosómica 28S S35, mitocondrial	2.101.680	1.357.920	1.555.814	1.555.814	23,04
NP	Purina nucleósido fosforilasa	60.775	67.168		63.971	7,07
PARP4	Poli [ADP-ribosa] polimerasa 4	476.849	523.643	243.332	476.849	36,22
PFDN1	Subunidad de prefoldina 1	8.881.373	8.377.838	8.781.079	8.781.079	3,07
PRDX6	Peroxirredoxina 6	1.062.048	950.200	1.192.875	1.062.048	11,37
PSMC3	Subunidad 6 reguladora de proteasa 26S	1.095.695	1.022.209	1.209.724	1.095.695	8,52
SSRP1	Subunidad SSRP1 de complejo FACT	239.173	184.576	204.962	204.962	13,16
THOC1	Subunidad 1 de complejo THO	397.408	278.736	357.637	357.637	17,53
TPR	Nucleoproteína TPR	1.022.450	713.318	1.025.854	1.022.450	19,5
UQCRC1	Subunidad de complejo de citocromo b-c1	22.974.646	17.376.010	22.886.339	22.886.339	15,22
VIM	Vimentina					

**Tabla Complementaria 4c (continuación):**

Experimento inverso (PrEST pesado)		Número de copias 2		Número de copias 3		Mediana	DTR (%)	DTR (%)
Número de copias 1	Número de copias 2	Número de copias 2	Número de copias 3	Número de copias 2	Número de copias 3	Mediana	DTR (%)	DTR (%)
152.131	6.745.143	5.350.929	152.131	18,11763	58,96924	152.131	11,84	58,96924
5.837.904	476.434	589.556	5.837.904	33,33	100,8334	5.837.904	33,33	100,8334
290.925	4.139.892	3.537.849	290.925	11,09	10,8856	476.434	11,09	10,8856
117.915	293.058	7.618	117.915	60,27	31,61052	3.838.870	60,27	31,61052
12.753	7.618	299.044	12.753	35,65	0,037884	205.487	35,65	0,037884
299.044	5.276.677	3.157.087	299.044	23,61427	23,61427	10.185	23,61427	23,61427
2.484.899	1.531.715	1.870.428	2.484.899	40,03	28,26456	299.044	40,03	28,26456
1.529.120	395.520	308.967	1.529.120	11,94	25,24945	3.157.087	11,94	25,24945
1.028.296	11.287.790	7.120.698	1.028.296	7,99	63,69902	1.531.715	7,99	63,69902
2.052.605	11.287.790	7.120.698	2.052.605	32,01	1,142762	395.520	32,01	1,142762
252.123	162.252	162.252	252.123	25,89	8,240033	2.052.605	25,89	8,240033
234.988	149.148	149.148	234.988	40,06	67,66706	9.204.244	40,06	67,66706
262.710	1.785.222	1.476.153	262.710	24,78	3,715642	9.204.244	24,78	3,715642
1.071.595	413.808	413.808	1.071.595	103,5509	103,5509	2.052.605	103,5509	103,5509
771.578	9.447.065	6.349.191	771.578	33,3868	33,3868	9.204.244	33,3868	33,3868
10.583.02	2.514.900	5.692.641	10.583.02	24,92	5,166992	9.447.065	24,92	5,166992
1.374.278	911.155	747.051	1.374.278	70,06	57,44122	2.514.900	70,06	57,44122
1.026.573	315.936	303.992	1.026.573	15,70	13,00441	911.155	15,70	13,00441
506.483	1.312.073	888.889	506.483	35,94	75,84125	679.026	35,94	75,84125
318.235	25.128.776	21.745.625	318.235	2,45	8,755396	315.936	2,45	8,755396
20.706.797	25.128.776	21.745.625	20.706.797	27,19	5,138108	1.100.481	27,19	5,138108
				10,26	3,614478	21.745.625	10,26	3,614478

*Quantificación absoluta en experimentos individuales* – También se deseó desarrollar una variación de la estrategia de SILAC-PrEST para cuantificar dianas proteicas individuales. En este caso, las dos etapas experimentales implicadas en la cuantificación de proteínas absolutas pueden colapsarse en una sola como se perfila esquemáticamente en la Figura 6A. Se mezcla una cantidad conocida con precisión del marcador de solubilidad ABP en lisado celular junto con el PrEST marcado. El análisis de CL-EM/EM de la muestra proporciona después relaciones de SILAC de marcador de solubilidad de ABP ligero con respecto a péptidos de ABP PrEST marcados. Estas relaciones cuantifican con precisión la cantidad de PrEST que se usó. Los mismos datos de CL EM también contienen las relaciones de péptidos de PrEST marcados con respecto al homólogo de proteína endógena no marcada. Juntas, estas relaciones cuantifican la cantidad absoluta de proteína endógena en un único experimento. Obsérvese que no se requiere marcaje de triple SILAC en este enfoque debido a que las relaciones se determinan frente a regiones diferentes de la construcción de PrEST, concretamente la región marcadora de solubilidad de ABP común (para cuantificar el PrEST) y la región de PrEST específica de proteína (para cuantificar la proteína endógena).

Este método individual para cuantificación se realizó para tres proteínas HeLa diferentes en las que el lisado celular marcado con SILAC y ABP marcada con SILAC se cuantificó frente a los PrEST no marcados. Como se muestra en la Figura 6C, se obtuvieron valores uniformes en estas mediciones basándose en experimentos por triplicado. Los niveles absolutos generalmente coincidieron bien con los números de copias determinados de forma independiente en el experimento múltiple de PrEST - SILAC descrito anteriormente (la diferencia máxima entre las medias de 40 %), validando ambos enfoques.

*Ensayo inmunoabsorbente ligado a enzimas* - ELISA es un método convencional en la investigación bioquímica para determinar cantidades absolutas o al menos para determinar de forma reproducible niveles de proteínas. Se comparó por lo tanto el método de SILAC - PrEST con esta tecnología establecida. Cuando se realiza el ensayo de ELISA para estratífina (14-3-3  $\sigma$ ) en condiciones típicas, lisado celular filtrado y solución salina tamponada con fosfato (PBS) como se recomienda por el fabricante, el ELISA registró menos del 20 % de la cantidad cuantificada por EM. (Obsérvese que no hay ninguna interferencia por las isoformas 14-3-3 porque estos péptidos son diferentes). La recomendación del fabricante fue que PBS no podría solubilizar el sedimento. La solubilidad se aumentó añadiendo el detergente no iónico NP-40, que fue capaz de disolver la mayoría del sedimento de muestra. Añadiendo una baja concentración de dodecilsulfato sódico (SDS), un detergente aniónico mejoró adicionalmente la cantidad de proteína medida significativamente aumentada (Figura 8B). Las cantidades absolutas aún se subestimaron el doble en comparación con el análisis de espectrometría de masas, supuestamente debido a que el protocolo de FASP permite la solubilización completa mediante el uso de dodecilsulfato sódico al 4 %.

También se investigaron los niveles del factor de transcripción y protooncogén FOS por ELISA, la proteína de menor abundancia cuantificada en la mezcla. Aquí la solubilización no parecía ser un problema y se recibió un excelente acuerdo entre los valores cuantitativos determinados por EM y por ELISA usando diferentes condiciones de tampón (Figure 8A).

### 40 Ejemplo 3: cuantificación absoluta usando PrEST de ratón

*Procedimiento experimental* – Los PrEST de ratón fusionados con un marcador de His N terminal se expresaron en una cepa de *E. coli* auxotrófica usando un medio mínimo, complementado con lisina  $^{13}\text{C}_6$   $^{15}\text{N}_2$  -Lisina (Lys8) y  $^{13}\text{C}_6$   $^{15}\text{N}_4$  -Arginina (Arg10) marcadas con isótopos (Cambridge Isotopes Laboratories) para obtener proteínas marcadas “pesadas”. Las bacterias se recogieron por centrifugación, se lisaron en cloruro de guanidinio 7 M,  $\text{Na}_2\text{HPO}_4$  47 mM,  $\text{NaH}_2\text{PO}_4$  2,65 mM, Tris HCl 10 mM, NaCl 300 mM, beta-mercaptoetanol 10 mM, pH 8,0 y los PrEST fusionados con His se enriquecieron en una columna Cobalt Talon (Clontech) y se eluyeron en urea 6 M,  $\text{NaH}_2\text{PO}_4$  50 mM, NaCl 100 mM, ácido acético 30 mM, acetato Na 70 mM pH 5 (29).

Se extrajeron muestras sanguíneas de ratones en tubos que contenían heparina. La sangre se centrifugó dos veces a 70 g y cada vez se conservó el sobrenadante, el plasma rico en plaquetas (PRP). Se añadieron apirasa y prostaciclina ( $\text{PGI}_2$ ) al PRP para inhibir la agregación plaquetaria. La muestra se centrifugó y el sedimento se lavó dos veces con 1 ml de tampón de Tyrode (sin  $\text{Ca}^{2+}$ , que contenía BSA, apirasa y  $\text{PGI}_2$ ). Con el tiempo el sedimento se resuspendió en 300-400  $\mu\text{l}$  de tampón de Tyrode y se incubó durante 30 min a 37 °C. Se realizó un análisis hematológico convencional usando el Hemavet 950 (Drew Scientific Inc.) para contar plaquetas.

Las plaquetas aisladas se lisaron en SDS 4 %, Tris 100 mM pH 8,5, DTT 100 mM, se hirvieron durante 5 min a 95 °C y se añadieron los PrEST purificados al lisado en la cantidad apropiada. Las muestras se prepararon de acuerdo con el método de FASP previamente descrito (30). Los péptidos se recogieron por centrifugación y se eluyeron con agua. Los péptidos se desalaron en stages tips empore C18 y se eluyeron en tampón B (acetonitrilo 80 %, ácido acético 0,5 %), se retiró disolvente orgánico por uso de speed-vac y la muestra se resolvió en A\* (acetonitrilo 2 %, ácido acético 0,5 %). Los péptidos se cargaron sin prefraccionamiento en una columna de 20 cm empaquetada de forma interna (75  $\mu\text{m}$  de diámetro interno) empaquetada con resina C18 de 1,8  $\mu\text{m}$  (Dr. Maisch GmbH) y se separó usando un EASY-nLC 1000 (Thermo Fisher Scientific) en un gradiente de tampón B de 2-25 % de 200 min. Los péptidos separados se pulverizaron mediante una fuente de iones de nanoelectropulverización (Proxeon Biosystems) a un espectrómetro de masas Q Exactive (Thermo Fisher Scientific). El espectrómetro de masas adquirió exploraciones

de sondeo y los 10 iones más abundantes superiores se fragmentaron secuencialmente con disociación de colisión de mayor energía y exploraciones de EM/EM adquiridas. Los datos sin procesar se analizaron usando el software Max Quant como se describe en el Ejemplo 1 excepto que los datos se buscaron en la base de datos IPI de ratón versión 3.68 que contenía 56.743 entradas.

5 **Resultados-** Para ampliar adicionalmente el enfoque a otras especies se diseñaron PrEST que se dirigían a proteínas de ratón. Se diseñaron PrEST para abarcar una región de 125-200 aminoácidos, produciendo muchos péptidos trípticos e incluyendo numerosos péptidos que se observaron en el espectrómetro de masas en mediciones previas. Para cada proteína diana se diseñaron dos PrEST para abarcar diferentes regiones de las proteínas y para asegurar la precisión de cuantificación. Se diseñaron PrEST para medir los niveles de expresión de integrina beta 3 y sus coactivadores talina 1 y kindlina 3 en plaquetas de ratón. La activación de la integrina heterodimérica  $\alpha\text{IIb}\beta\text{3}$  (que desplaza de un estado de baja afinidad a un estado de alta afinidad) desempeña un papel esencial en la adhesión y agregación de plaquetas (31). Los ratones deficientes en kindlina 3 padecen hemorragia grave y mueren en un periodo de varios días. Se determinaron los niveles de expresión de integrina beta 3, talina 1 y kindlina 3 en ratones de tipo silvestre (Kind3<sup>+/+</sup>), Kind3<sup>+/-</sup>, Kind3<sup>n/n</sup> y Kind3<sup>n/-</sup>. 'n' indica una inserción de un casete de neomicina en un intrón del gen, lo que afecta al corte y empalme de kindlina 3. Para dilucidar adicionalmente la funcionalidad de activación de integrina se deseó medir la estequiometría de integrina beta 3, talina 1 y kindlina 3 en los ratones de tipo silvestre.

20 La integrina beta 3 y sus coactivadores son proteínas altamente abundantes en plaquetas e Itgb3 tiene en promedio 300.000 copias por célula, mientras que sus coactivadores talina 1 tiene 470.000 copias y kindlina 3 tiene en promedio 430.000 copias por plaqueta, (Tabla 3, Figura 9a). Se midieron números de copias de las proteínas diana en los diferentes ratones por duplicado y usando dos PrEST diferentes. La diferencia entre muestras de plaquetas fue en promedio 20 %, mientras que la diferencia entre PrEST es del 22 %. Para el cálculo de kindlina 3 solamente se consideró un PrEST ya que este se dirige a la región de interés biológico, el extremo C terminal de kindlina 3 interacciona con la cola citoplasmática de integrina beta 3.

30 Además de las copias por célula también se observó la reducción del nivel de expresión de kindlina 3 en los diferentes desactivados (Figura 9b). En comparación con los ratones de tipo silvestre la kindlina 3 se redujo como se esperaba hasta el 50 % en los ratones Kind3<sup>+/-</sup>, hasta el 15 % en los ratones Kind3<sup>n/n</sup> y hasta el 6 % en los ratones Kind3<sup>n/-</sup> y la tendencia está de acuerdo con las observaciones de Moser *et al.* (32).

**Tabla 3:** Números de copias por plaqueta. Las cantidades absolutas de las proteínas de interés se midieron cada una usando dos PrEST en dos muestras de ratones.

	Integrina beta 3	Talina 1	Kindlina 3	%
Kind3 <sup>+/+</sup>	345.000	531.000	433.000	100
Kind3 <sup>+/-</sup>	306.000	445.000	242.000	55
Kind3 <sup>n/n</sup>	313.000	490.000	68.000	15
Kind3 <sup>n/-</sup>	268.000	409.000	26.000	6

35 Usando la cantidad absoluta se determinó que la estequiometría de las tres proteínas (Tabla 2) en los ratones de tipo silvestre era de 1:1,5:1,3 y esta información de estequiometría ayuda a entender adicionalmente la unión de coactivadores y la activación de integrinas.

40 **Tabla 4:** Estequiometría de la proteína calculada en ratones de tipo silvestre usando los niveles de expresión absolutos.

	Integrina beta 3	Talina 1	Kindlina 3
Número de copias	345.000	531.000	433.000
Estequiometría	1	1,5	1,3

#### Referencias adicionales

- 45 1. Aebersold, R., y Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* 422, 198-207.  
 2. Cravatt, B. F., Simon, G. M., y Yates, J. R., 3rd (2007) The biological impact of mass spectrometry-based proteomics. *Nature* 450, 991-1000.  
 3. Gstaiger, M., y Aebersold, R. (2009) Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet* 10, 617-627.  
 50 4. Ong, S. E., y Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nature chemical biology* 1, 252-262.  
 5. Bachi, A., y Bonaldi, T. (2008) Quantitative proteomics as a new piece of the systems biology puzzle. *J Proteomics* 71,357-367.  
 6. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., y Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 389, 1017-1031.  
 55 7. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., y Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376- 386.

8. Mann, M. (2006) Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol* 7, 952-958.
9. Geiger, T., Wisniewski, J. R., Cox, J., Zanivan, S., Kruger, M., Ishihama, Y., y Mann, M. (2011) Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nature protocols* 6, 147-157.
- 5 10. Brun, V., Masselon, C., Garin, J., y Dupuis, A. (2009) Isotope dilution strategies for absolute quantitative proteomics. *J Proteomics* 72, 740-749.
11. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., y Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6940-6945.
- 10 12. Beynon, R. J., Doherty, M. K., Pratt, J. M., y Gaskell, S. J. (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nature methods* 2, 587-589.
13. Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., Gaskell, S. J., y Beynon, R. J. (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nature protocols* 1, 1029-1043.
- 15 14. Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., y Garin, J. (2007) Isotope-labeled protein standards: toward absolute quantitative proteomics. *Mol Cell Proteomics* 6, 2139-2149.
15. Hanke, S., Besir, H., Oesterhelt, D., y Mann, M. (2008) Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *Journal of proteome research* 7, 1118-1130.
16. Singh, S., Springer, M., Steen, J., Kirschner, M. W., y Steen, H. (2009) FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. *Journal of proteome research* 8, 2201-2210.
- 20 17. Kuster, B., Schirle, M., Mallick, P., y Aebersold, R. (2005) Scoring proteomes with proteotypic peptide probes. *Nat Rev Mol Cell Biol* 6, 577-583.
18. Berglund, L., Bjorling, E., Jonasson, K., Rockberg, J., Fagerberg, L., Al-Khalili Szgyarto, C., Sivertsson, A., y Uhlen, M. (2008) A whole-genome bioinformatics approach to selection of antigens for systematic antibody generation. *Proteomics* 8, 2832-2839.
- 25 19. Larsson, M., Graslund, S., Yuan, L., Brundell, E., Uhlen, M., Hoog, C., y Stahl, S. (2000) High throughput protein expression of cDNA products as a tool in functional genomics. *Journal of biotechnology* 80, 143-157.
20. Agaton, C., Galli, J., Hoiden Guthenberg, I., Janzon, L., Hansson, M., Asplund, A., Brundell, E., Lindberg, S., Ruthberg, I., Wester, K., Wurtz, D., Hoog, C., Lundberg, J., Stahl, S., Ponten, F., y Uhlen, M. (2003) Affinity proteomics for systematic protein profiling of chromosome 21 gene products in human tissues. *Mol Cell Proteomics* 2, 405-414.
- 30 21. Li, M. Z., y Elledge, S. J. (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nature methods* 4, 251-256.
22. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* 41, 207-234.
- 35 23. Wisniewski, J. R., Zougman, A., Naga raj, N., y Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nature methods* 6, 359-362.
24. Wisniewski, J. R., Zougman, A., y Mann, M. (2009) Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. *Journal of proteome research* 8, 5674-5678.
- 40 25. Rappsilber, J., Ishihama, Y., y Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical chemistry* 75, 663-670.
26. Geiger, T., Cox, J., y Mann, M. (2010) Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol Cell Proteomics* 9, 2252-2261.
- 45 27. Cox, J., y Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* 26, 1367-1372.
28. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., y Mann, M. (2011) Andromeda - a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research*.
- 50 29. Tegel, H., Steen, J., Konrad, A., Nikdin, H., Pettersson, K., Stenvall, M., Tourle, S., Wrethagen, U., Xu, L., Yderland, L., Uhlen, M., Hober, S., y Ottosson, J. (2009) High-throughput protein production-lessons from scaling up from 10 to 288 recombinant proteins per week. *Biotechnol J* 4, 51-57.
30. Wisniewski, J. R., Zougman, A., Nagaraj, N., y Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat Methods* 6, 359-362.
- 55 31. Moser, M., Legate, K. R., Zent, R., y Fassler, R. (2009) The tail of integrins, talin, and kindlins. *Science* 324, 895-899.
32. Moser, M., Nieswandt, B., Ussar, S., Pozgajova, M., y Fassler, R. (2008) Kindlin-3 is essential for integrin activation and platelet aggregation. *Nat Med* 14, 325-330.
33. Matic, I., Jaffray, E. G., Oxenham, S. K., Groves, M. J., Barratt, C. L., Tauro, S., Stanley-Wall, N. R., y Hay, R. T. (2011) Absolute SILAC-compatible expression strain allows Sumo-2 copy number determination in clinical samples. *J Proteome Res* 10, 4869-4875.
- 60 34. Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., Damoc, E., Denisov, E., Lange, O., Remes, P., Taylor, D., Splendore, M., Wouters, E. R., Senko, M., Makarov, A., Mann, M., y Horning, S. (2009) A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol Cell Proteomics* 8, 2759-2769.
- 65

A continuación hay un listado de realizaciones desglosado de la presente divulgación.

1. Un método para determinar la cantidad absoluta de un polipéptido diana en una muestra, comprendiendo dicho método las siguientes etapas:

- 5 (a) añadir  
 (aa) un polipéptido de fusión a dicha muestra, comprendiendo dicho polipéptido de fusión (i) al menos una secuencia marcadora y (ii) una subsecuencia del polipéptido diana; y  
 (ab) una cantidad absoluta conocida de un polipéptido marcador que comprende o consiste en dicha secuencia marcadora de acuerdo con (aa)  
 10 a dicha muestra, en el que dicho polipéptido de fusión por un lado tiene masa alterada en comparación con dicho polipéptido diana y dicho polipéptido marcador por otro lado, por ejemplo, dicho polipéptido de fusión por un lado y dicho polipéptido diana y dicho polipéptido marcador por otro lado están marcados con isotópicos de forma diferente;  
 (b) realizar una digestión proteolítica de la mezcla obtenida en la etapa (a);  
 15 (c) someter el resultado de la digestión proteolítica de la etapa (b), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y  
 (d) determinar la cantidad absoluta de dicho polipéptido diana de (i) las intensidades pico en el espectro de masas adquirido en la etapa (c) de dicho polipéptido de fusión, dicho polipéptido marcador y dicho polipéptido diana y (ii) dicha cantidad absoluta conocida de dicho polipéptido marcador.

20 2. Un método para crear un patrón cuantitativo, comprendiendo dicho método las siguientes etapas:

- (a) proporcionar uno o una pluralidad de polipéptidos de fusión, comprendiendo el polipéptido de fusión o cada uno de dichos polipéptidos de fusión, respectivamente, (i) al menos una secuencia marcadora y (ii) una subsecuencia de un polipéptido diana para determinar de forma cuantitativa, en el que, en la medida en que se proporciona dicha pluralidad de polipéptidos de fusión, todos los polipéptidos de fusión comparten al menos una secuencia marcadora, obteniendo de este modo el patrón;  
 25 (b) determinar las cantidades absolutas de dichos polipéptidos de fusión  
 (ba) añadiendo al polipéptido de fusión o a dichos polipéptidos de fusión a la vez, respectivamente, una cantidad conocida de un polipéptido marcador que comprende o consiste en la secuencia marcadora comprendida en el polipéptido de fusión o compartida entre los polipéptidos de fusión, respectivamente, de acuerdo con (a), en el que dicho polipéptido de fusión tiene masa alterada en comparación con dicho polipéptido marcador, por ejemplo, dicho polipéptido de fusión y dicho polipéptido marcador están marcados con isótopos de forma diferente,  
 30 (bb) realizando digestión proteolítica de la mezcla de un polipéptido de fusión y dicho polipéptido marcador obtenido en la etapa (ba);  
 (bc) sometiendo el resultado de la digestión proteolítica de la etapa (bb), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y  
 (bd) determinando la cantidad absoluta de dicho polipéptido de fusión de (i) las intensidades pico en el espectro de masas de polipéptido de fusión y polipéptido marcador y (ii) dicha cantidad conocida de dicho polipéptido marcador,  
 35 obteniendo de este modo la cantidad absoluta del polipéptido de fusión o de uno de dicha pluralidad de polipéptidos de fusión a la vez, respectivamente.

45 3. Un método para determinar la cantidad absoluta de uno o más polipéptidos diana en una muestra, comprendiendo dicho método las siguientes etapas:

- (a) opcionalmente realizar el método del artículo 2;  
 (b) añadir el patrón cuantitativo como se ha definido en el artículo 2 (a) a dicha muestra;  
 50 (c) realizar digestión proteolítica de la mezcla obtenida en la etapa (b);  
 (d) someter el resultado de digestión proteolítica de la etapa (c), opcionalmente después de cromatografía, a análisis espectrométrico de masas; y  
 (e) determinar las cantidades absolutas del polipéptido o los polipéptidos diana de (i) las intensidades picos en el espectro de masas adquirido en la etapa (d) de polipéptido o polipéptidos de fusión y polipéptido o polipéptidos diana y (ii) la cantidad o las cantidades absolutas conocidas de dicho polipéptido o dichos polipéptidos de fusión,  
 55 en el que dicho polipéptido o dichos polipéptidos de fusión tienen masa alterada en comparación con dicho polipéptido o dichos polipéptidos diana, por ejemplo, dichos uno o más polipéptidos diana están marcados con isótopos de forma diferente en comparación con dichos polipéptidos de fusión.

60 4. El método de uno cualquiera de los artículos precedentes, en el que uno o dos marcadores están presentes en dicho polipéptido o dichos polipéptidos de fusión, seleccionándose dicho marcador o dichos marcadores de un marcador de purificación y un marcador de solubilidad.

65 5. El método de uno cualquiera de los artículos 1, 3 o 4, en el que dicha adición se efectúa antes de dicha digestión proteolítica.

6. El método de uno cualquiera de los artículos 2 a 5, en el que se usan entre dos y 500 polipéptidos de fusión.
7. El método de uno cualquiera de los artículos precedentes, en el que está presente un marcador de solubilidad en cada uno de dichos polipéptidos de fusión.
- 5 8. El método de uno cualquiera de los artículos precedentes, en el que dicha subsecuencia de un polipéptido  
 (a) consiste en de 15 a 205 aminoácidos;  
 (b) comprende un polipéptido proteotípico; y/o  
 (c) se selecciona para tener identidad de secuencia mínima con otras proteínas, excluye péptidos de señal y/o  
 10 excluye secuencias de regiones transmembrana.
9. Un polipéptido de fusión para la cuantificación de un polipéptido diana por espectroscopia de masas, en el que:  
 dicho polipéptido de fusión consiste en 35-455 restos de aminoácidos y comprende (i) una región diana, que es  
 15 un fragmento del polipéptido diana y (ii) una región marcadora, que no es un fragmento de polipéptido diana,  
 dicha región diana consiste en 15-205 restos de aminoácidos y comprende al menos dos regiones de  
 identificación;  
 dicha región marcadora consiste en 20-250 restos de aminoácidos y comprende al menos dos regiones de  
 identificación;  
 20 cada región de identificación tiene la estructura Y-Z-X<sub>4-28</sub>-Y-Z, en la que  
 todas las Y se seleccionan de uno de (i)-(iv), en los que (i) es R o K, (ii) es Y, F, W o L, (iii) es E y (iv) es D, y  
 cada X y cada Z son de forma independiente cualquier resto de aminoácido, siempre que las Z no sean P si las Y  
 se seleccionan de (i)-(iii); y cada región de identificación comprende al menos un resto de aminoácido que  
 comprende un isótopo pesado.
- 25 10. Un polipéptido de fusión de acuerdo con el artículo 9, en el que dicha región marcadora corresponde a un  
 marcador de solubilidad o un fragmento del mismo, seleccionándose dicho marcador de solubilidad de proteínas de  
 unión a maltosa (MBP), glutatión-S-transferasa (GST), tiorredoxina (Trx), sustancia de utilización de N (NusA),  
 modificador de ubiquitina pequeño (SUMO), un marcador potenciador de solubilidad (SET), una proteína formadora  
 30 de disulfuro C (DsbC), proteína de diecisiete kilodalton (Skp), proteína quinasa de fago T7 (T7PK), dominio B1 de  
 proteína G (GB1), dominio de repetición ZZ de IgG de proteína A (ZZ) y proteína de unión a albúmina (ABP).
11. Un polipéptido de fusión de acuerdo con el artículo 9 o 10, en el que dicha región marcadora consiste en o  
 comprende la secuencia expuesta en SEQ ID NO: 1.
- 35 12. Un polipéptido de fusión de acuerdo con uno cualquiera de los artículos 9 a 11, en el que las Y se seleccionan de  
 R y K.
13. El método de acuerdo con uno cualquiera de los artículos 1 a 8, en el que dicho polipéptido o dichos polipéptidos  
 40 de fusión son como se ha definido en uno cualquiera de los artículos 9 a 12.
14. Un kit que comprende:  
 (a) al menos un polipéptido de fusión de acuerdo con uno cualquiera de los artículos 9 a 12; y  
 (b) (i) un segundo polipéptido que comprende o consiste en la secuencia de aminoácidos de la región marcadora  
 45 como se ha definido en uno cualquiera de los artículos 9 a 12, estando dicho segundo polipéptido marcado  
 con isótopos de forma diferente en comparación con dicha región marcadora como se ha definido en uno  
 cualquiera de los artículos 9 a 12; y/o  
 (ii) una enzima proteolítica, tal como tripsina, quimotripsina, Lys-C, Glu-C o Asp-N.
15. Uso de un patrón cuantitativo como se ha definido en el artículo 2 o de un polipéptido de fusión de acuerdo con  
 50 uno cualquiera de los artículos 9 a 12 como una referencia en una cuantificación de polipéptido diana.

**Listado de secuencias**

- 55 <110> Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.  
 Atlas Antibodies AB
- <120> Patrón cuantitativo para espectrometría de masas de proteínas
- <130> T1588 EP/1
- 60 <150> EP 11 00 2794.3  
 <151> 04-04-2011
- <150> US 61/471.528
- 65 <151> 04-04-2011

ES 2 589 530 T3

<150> US 61/471.534  
 <151> 04-04-2011

<160> 1

5

<170> PatentIn versión 3.4

<210> 1

<211> 181

10

<212> PRT

<213> *streptococcus*

<400> 1

Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro  
 1 5 10 15

Arg Gly Ser His Met Ala Ser Leu Ala Glu Ala Lys Val Leu Ala Asn  
 20 25 30

Arg Glu Leu Asp Lys Tyr Gly Val Ser Asp Tyr His Lys Asn Leu Ile  
 35 40 45

Asn Asn Ala Lys Thr Val Glu Gly Val Lys Asp Leu Gln Ala Gln Val  
 50 55 60

Val Glu Ser Ala Lys Lys Ala Arg Ile Ser Glu Ala Thr Asp Gly Leu  
 65 70 75 80

Ser Asp Phe Leu Lys Ser Gln Thr Pro Ala Glu Asp Thr Val Lys Ser  
 85 90 95

Ile Glu Leu Ala Glu Ala Lys Val Leu Ala Asn Arg Glu Leu Asp Lys  
 100 105 110

Tyr Gly Val Ser Asp Tyr Tyr Lys Asn Leu Ile Asn Asn Ala Lys Thr  
 115 120 125

15

Val Glu Gly Val Lys Ala Leu Ile Asp Glu Ile Leu Ala Ala Leu Pro  
 130 135 140

Gly Thr Phe Ala His Tyr Gly Ser Ala Trp Ser His Pro Gln Phe Glu  
 145 150 155 160

Lys Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Ser Ala Trp Ser His  
 165 170 175

Pro Gln Phe Glu Lys  
 180

20

## REIVINDICACIONES

1. Un polipéptido de fusión para la cuantificación de un polipéptido diana por espectroscopia de masas, en el que:
  - 5 dicho polipéptido de fusión consiste en 35-455 restos de aminoácidos y comprende (i) una región diana, que es un fragmento del polipéptido diana y (ii) una región marcadora, que no es un fragmento de polipéptido diana, dicha región diana consiste en 15-205 restos de aminoácidos y comprende al menos dos regiones de identificación; dicha región diana consiste en 20-250 restos de aminoácidos y comprende al menos dos regiones de
    - 10 identificación; cada región de identificación tiene la estructura Y-Z-X<sub>4-28</sub>-Y-Z, todas las Y se seleccionan de uno de (i)-(iv), en los que (i) es R o K, (ii) es Y, F, W o L, (iii) es E y (iv) es D, y cada X y cada Z son de forma independiente cualquier resto de aminoácido, siempre que las Z no sean P si las Y se seleccionan de (i)-(iii);
    - 15 cada región de identificación comprende al menos un resto de aminoácido que comprende un isótopo pesado; dicha región marcadora consiste en proteína de unión a albúmina (ABP) o un fragmento de la misma; y el mismo resto Y puede constituir el extremo carboxílico de una primera región de identificación y el extremo amino de una segunda región de identificación.
  - 20 2. Un polipéptido de fusión de acuerdo con la reivindicación 1, en el que dicha región marcadora consiste en o comprende la secuencia expuesta en SEQ ID NO: 1.
  3. Un polipéptido de fusión de acuerdo con la reivindicación 1 o 2, en el que las Y se seleccionan de R y K.
  - 25 4. Un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 3, en el que dicha región diana consiste en de 20 a 150 restos de aminoácidos.
  5. Un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que dicha región marcadora consiste en de 40 a 150 restos de aminoácidos.
  - 30 6. Un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 5, que consiste en 80-300 restos de aminoácidos.
  7. Un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que el al menos un
    - 35 resto de aminoácido que comprende un isótopo pesado se selecciona de L-arginina-<sup>13</sup>C<sub>6</sub>, L-arginina-<sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>4</sub>, L-arginina-<sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>4</sub>D<sub>7</sub>, L-arginina-<sup>15</sup>N<sub>4</sub>D<sub>7</sub>, L-arginina-<sup>15</sup>N<sub>4</sub>, L-lisina-<sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>2</sub>, L-lisina-<sup>15</sup>N<sub>2</sub>, L-lisina-<sup>13</sup>C<sub>6</sub>, L-lisina-<sup>13</sup>C<sub>6</sub><sup>15</sup>N<sub>2</sub>D<sub>9</sub>, L-lisina-<sup>15</sup>N<sub>2</sub>D<sub>9</sub>, L-lisina-D<sub>4</sub>, L-metionina-<sup>13</sup>CD<sub>3</sub>, L-tirosina-<sup>13</sup>C<sub>9</sub>, L-tirosina-<sup>15</sup>N y L-tirosina-<sup>13</sup>C<sub>9</sub><sup>15</sup>N.
  - 40 8. Un polipéptido de fusión de acuerdo con la reivindicación 7, en el que todos los restos de arginina y lisina están marcados.
  9. Un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 8 que comprende además un marcador de purificación.
  - 45 10. Un polipéptido de fusión de acuerdo con la reivindicación 9, en el que el marcador de purificación se selecciona de un marcador de His, un marcador FLAG, un marcador SBP, un marcador myc y un marcador OneStrep.
  11. Un kit que comprende:
    - 50 (a) al menos un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 10; y
    - (b) (i) un segundo polipéptido que comprende o consiste en la secuencia de aminoácidos de la región marcadora como se ha definido en una cualquiera de las reivindicaciones 1 a 10, estando dicho segundo polipéptido marcado con isótopo de forma diferente en comparación con dicha región marcadora como se ha definido en una cualquiera de las reivindicaciones 1 a 10; y/o
    - 55 (ii) una enzima proteolítica, tal como tripsina, quimotripsina, Lys-C, Glu-C o Asp-N.
  12. Una fase sólida adecuada para digestión proteolítica, tal como un soporte sólido, columna o filtro, en el que se carga un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 10.
  - 60 13. Una columna en o sobre la que se dispone un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1-10 y una enzima proteolítica, tal como tripsina, quimotripsina, Lys-C, Glu-C o Asp-N.
  14. Uso de un polipéptido de fusión de acuerdo con una cualquiera de las reivindicaciones 1 a 10 como referencia en una cuantificación de polipéptido diana.
  - 65 15. Uso de acuerdo con la reivindicación 14, en el que la cuantificación de polipéptido diana se efectúa por

espectrometría de masas.

Figura 1

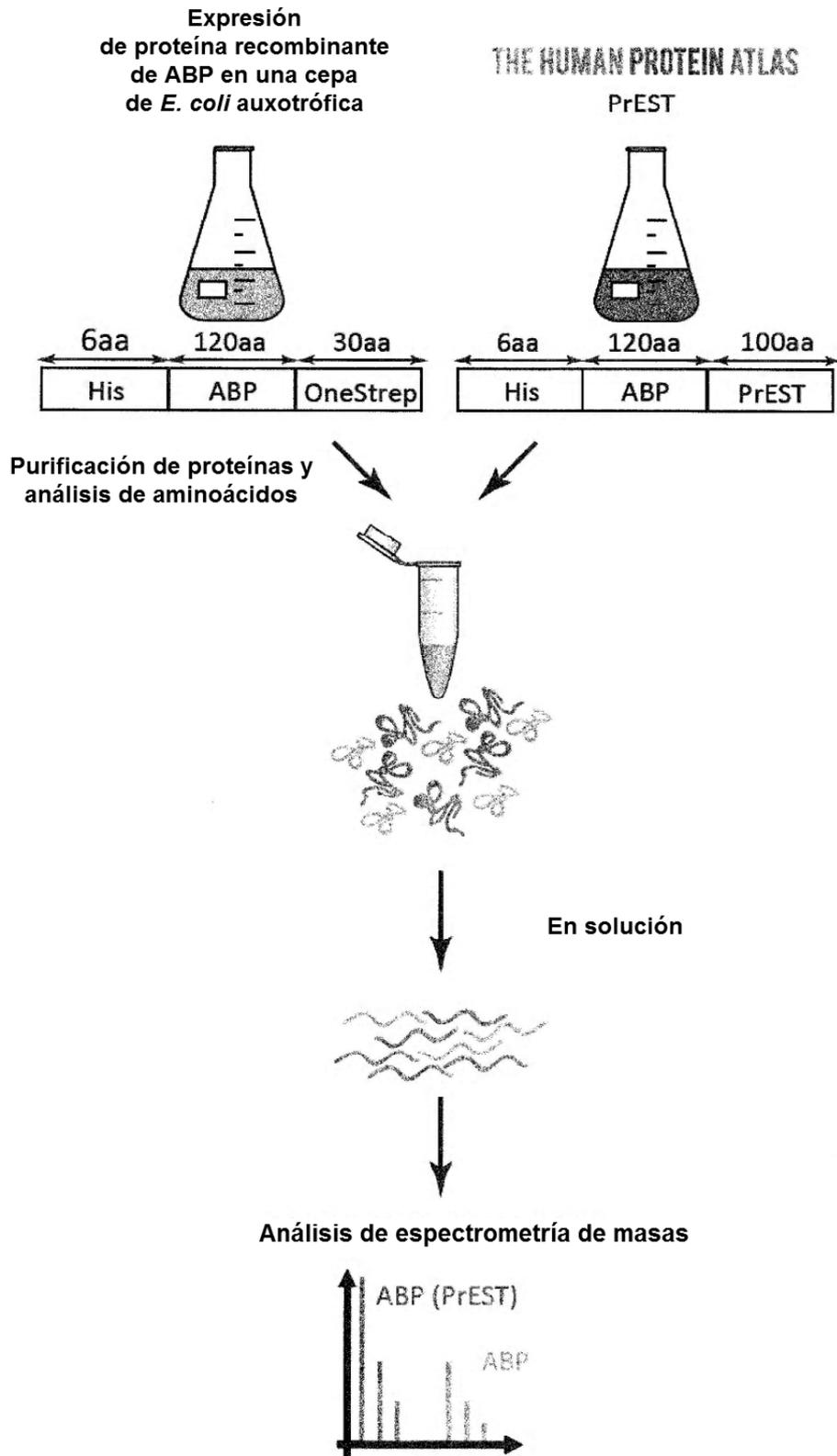


Figura 2a

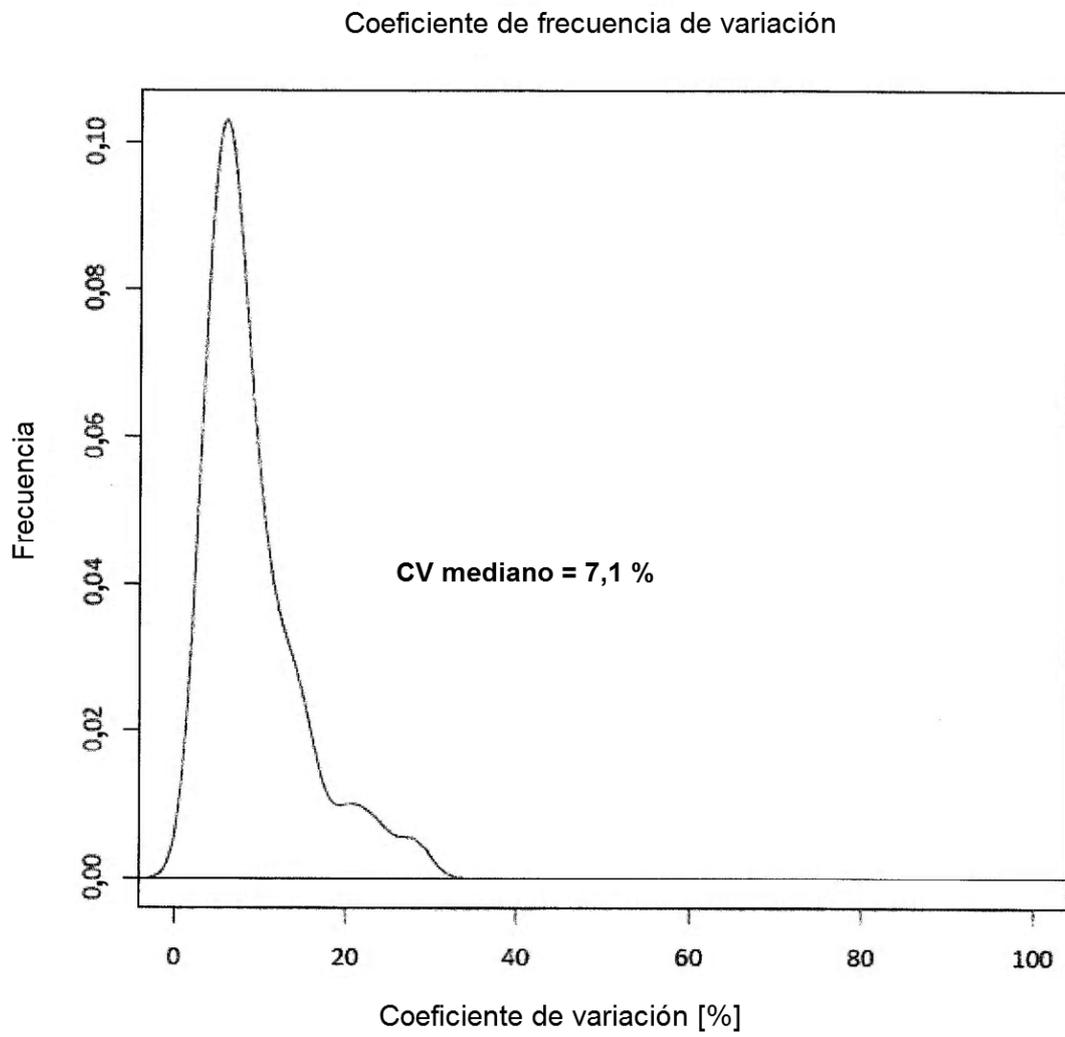
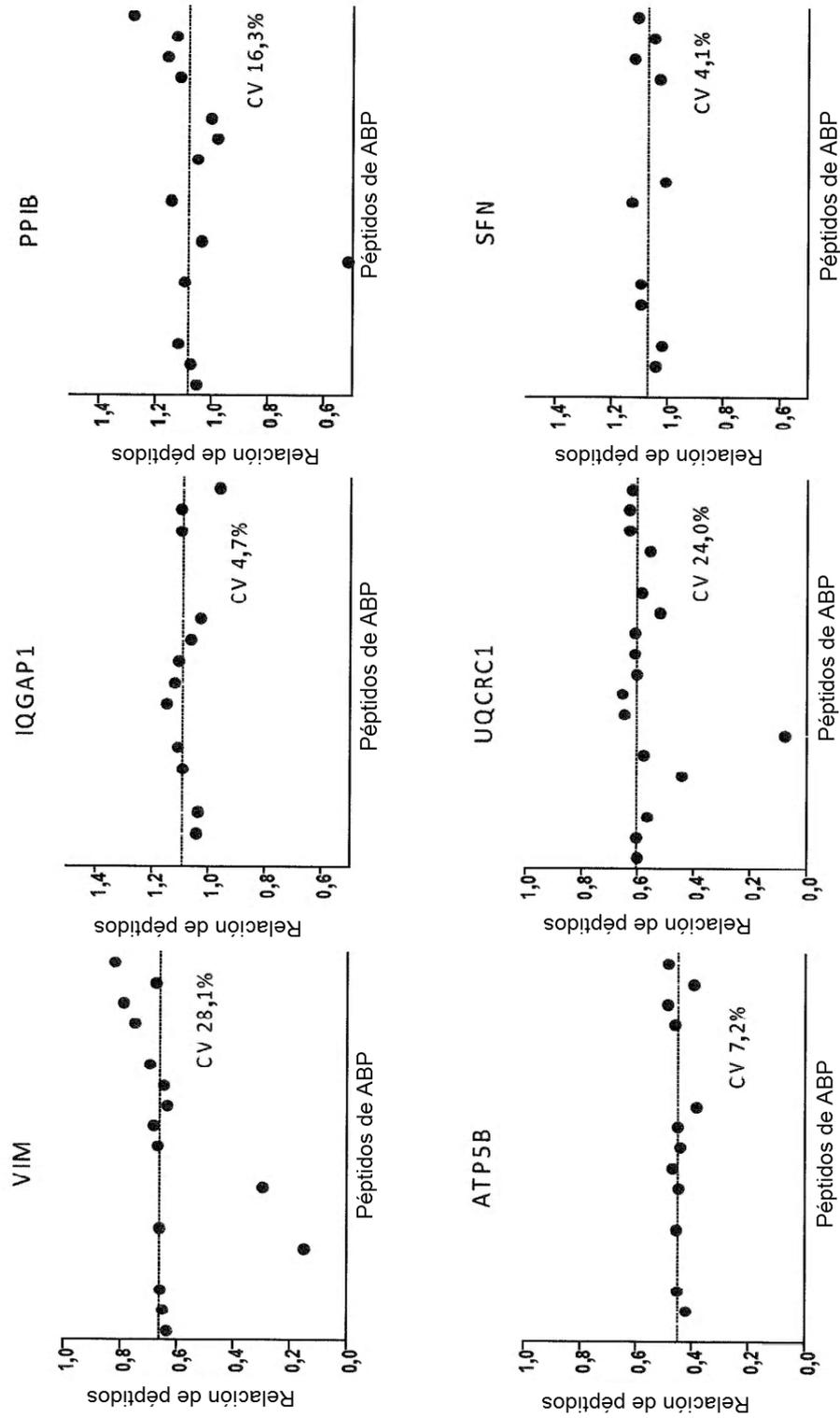


Figura 2b



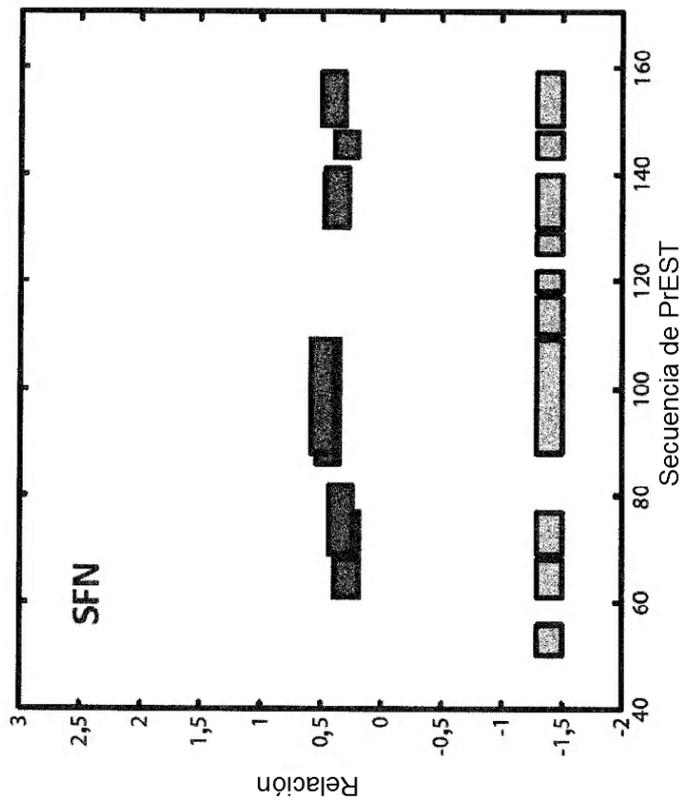
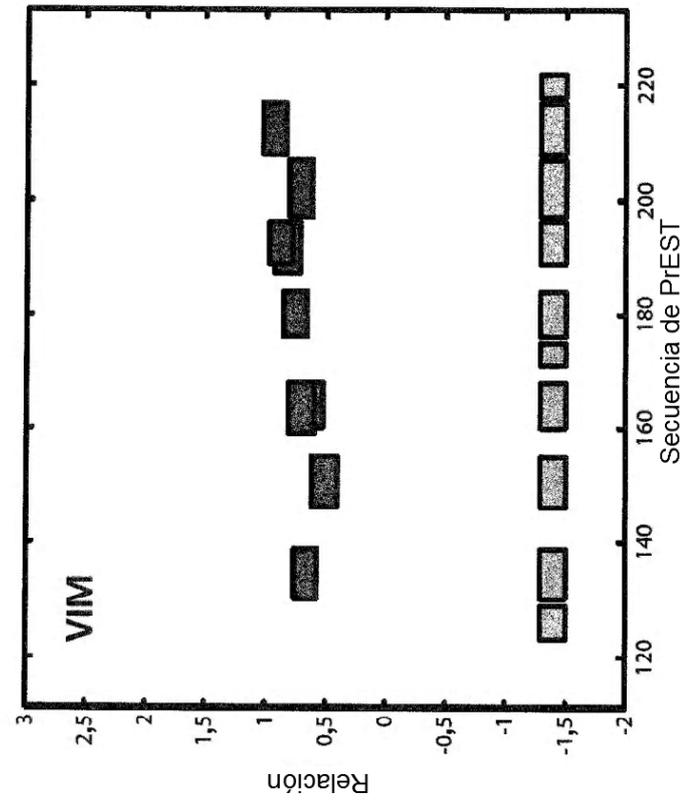


Figura 3

Figura 3 continuación

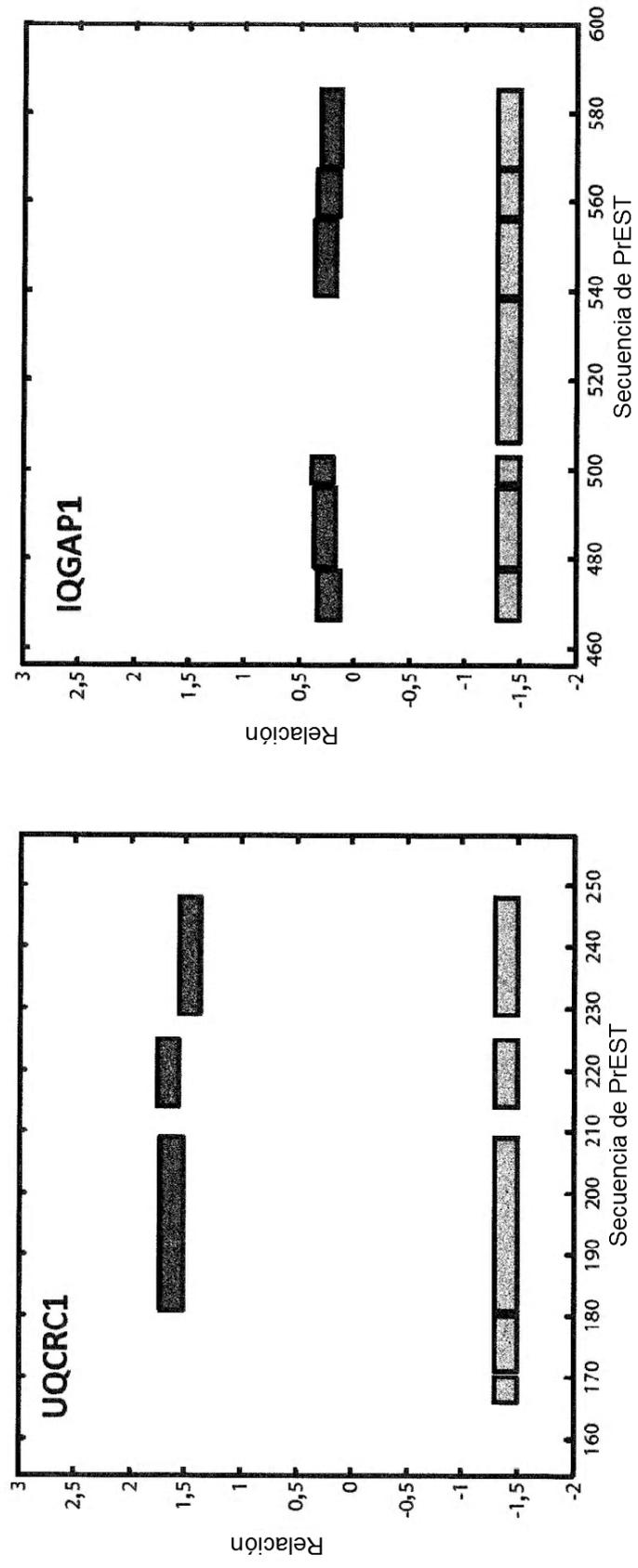


Figura 3 continuación

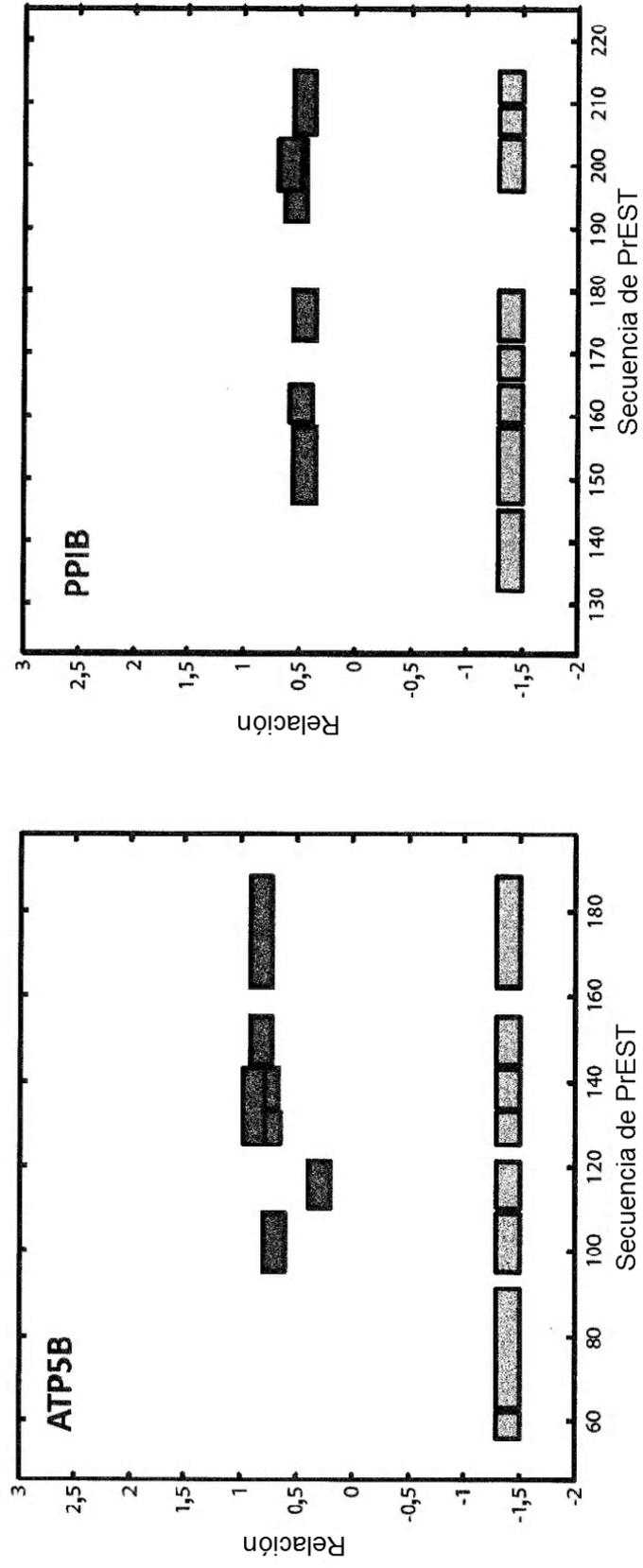


Figura 4

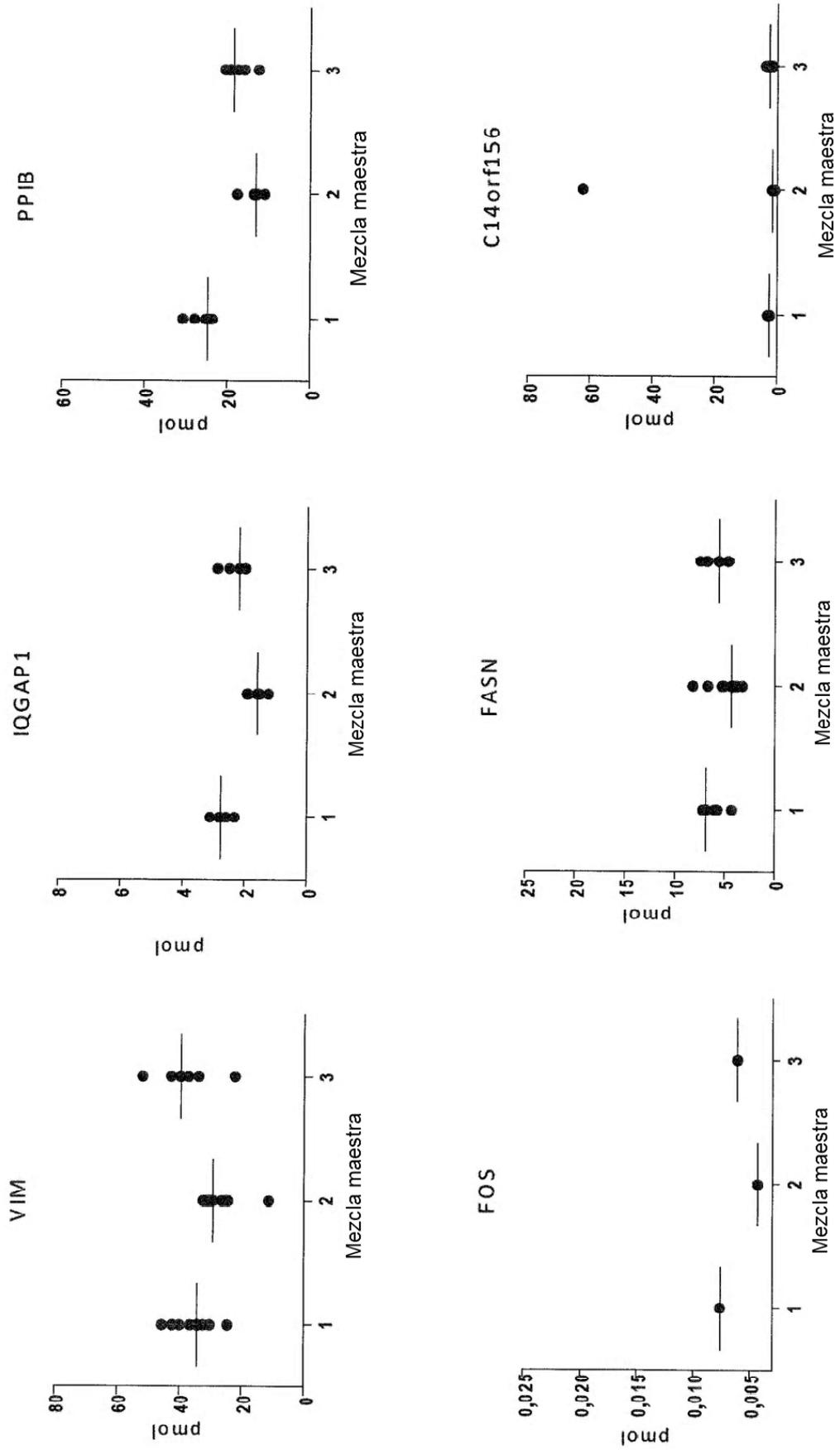


Figura 5

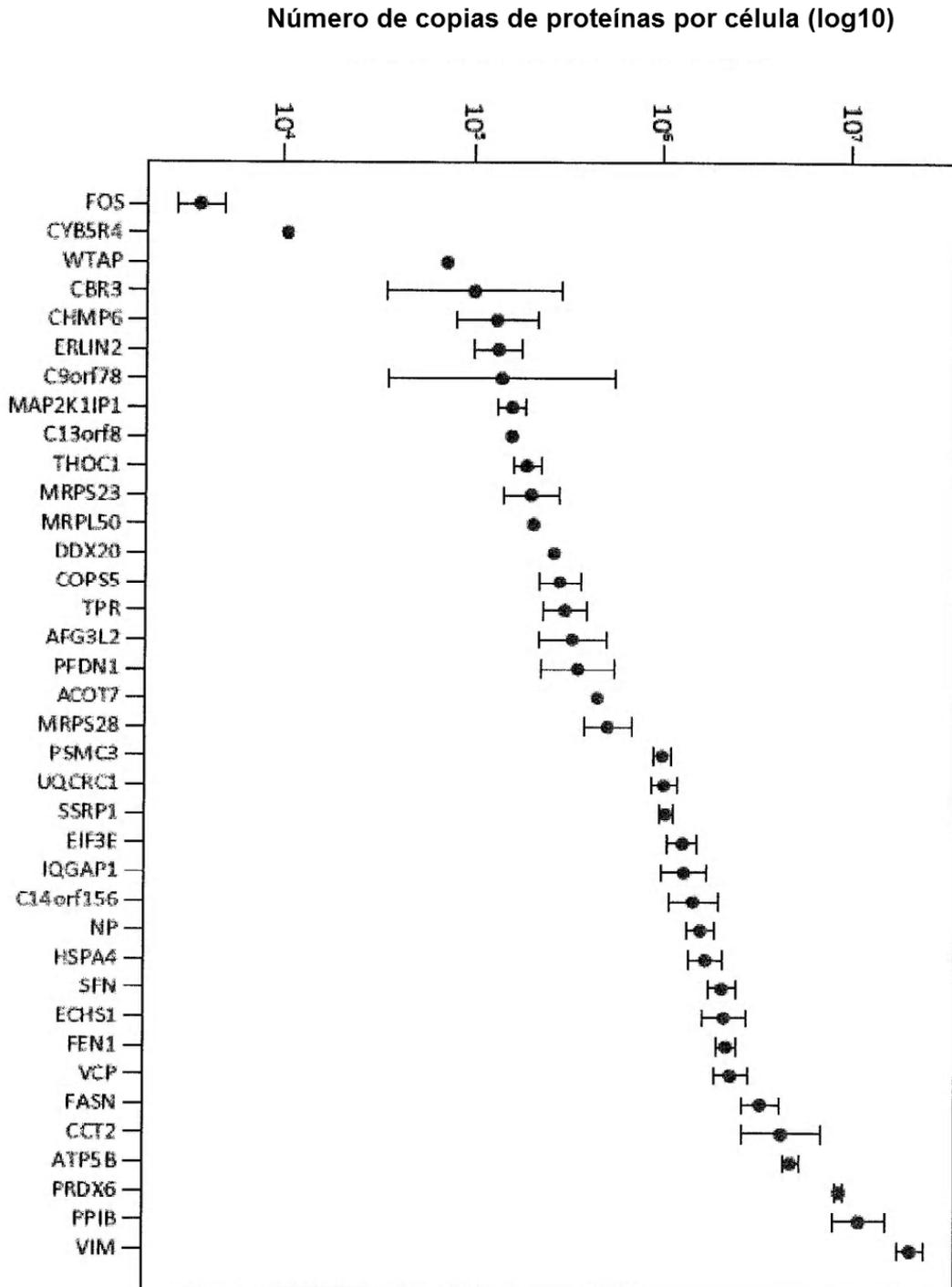


Figura 6a

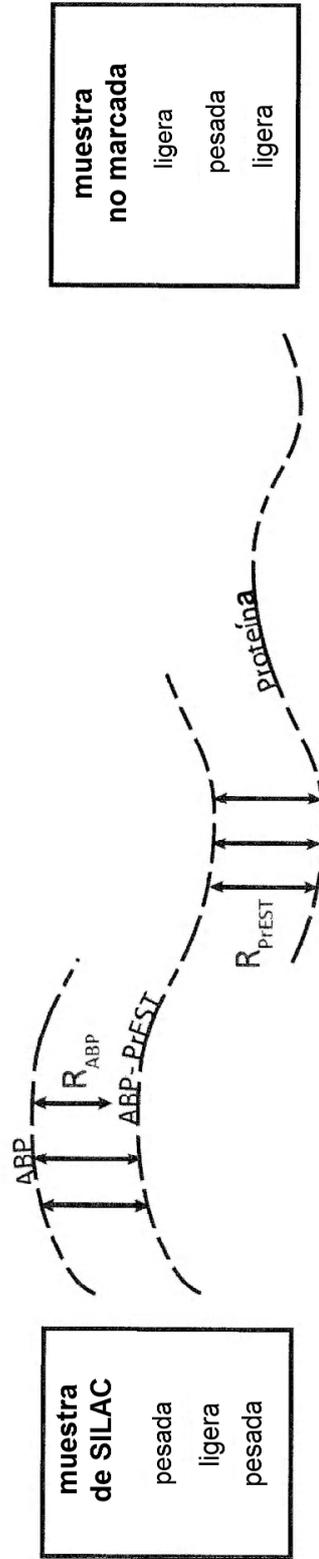


Figura 6b

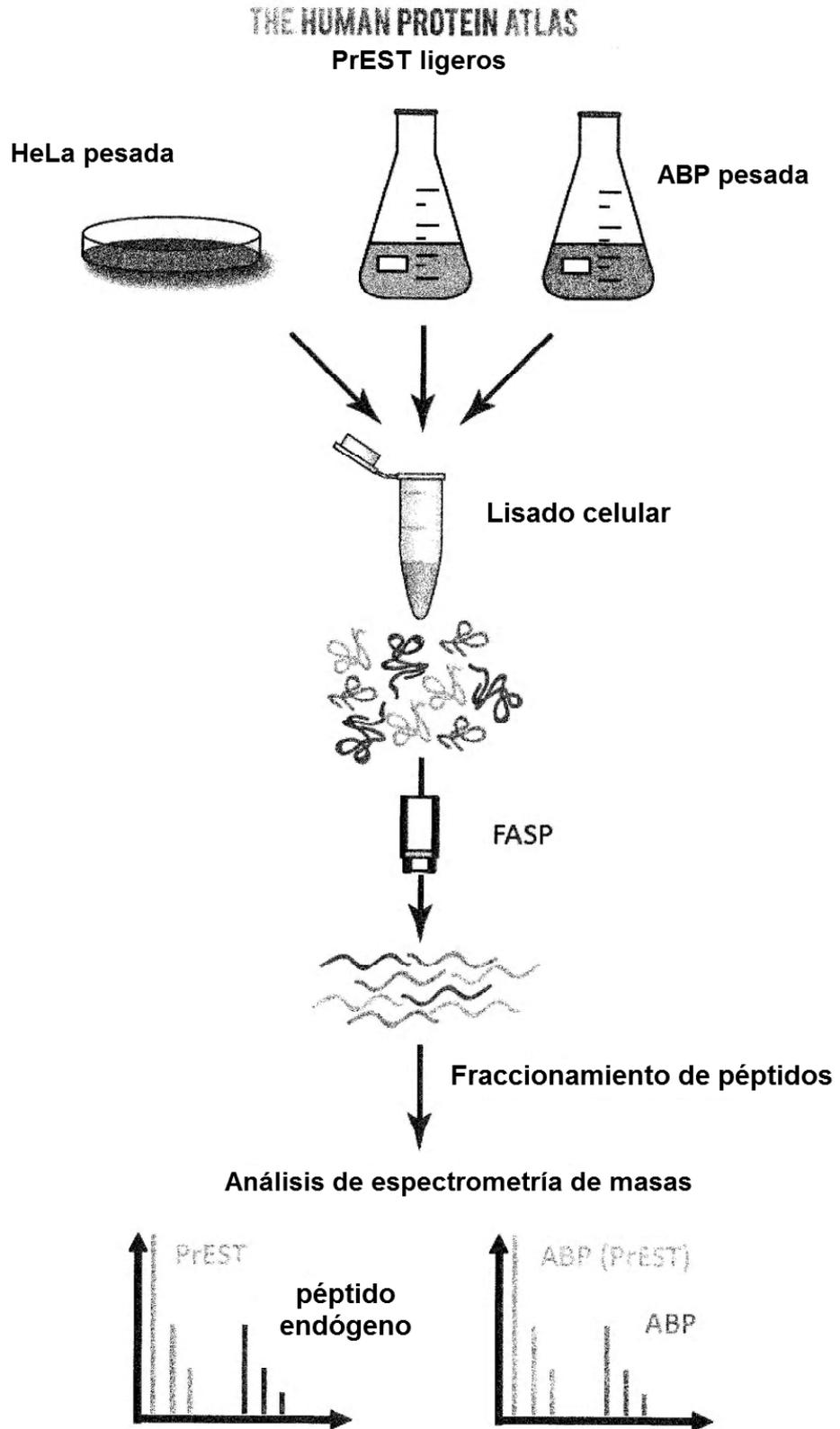


Figura 6c

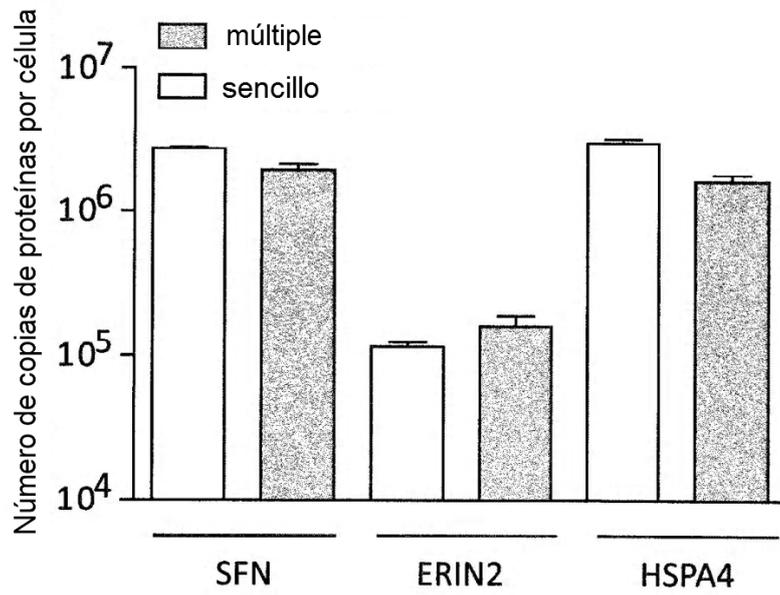
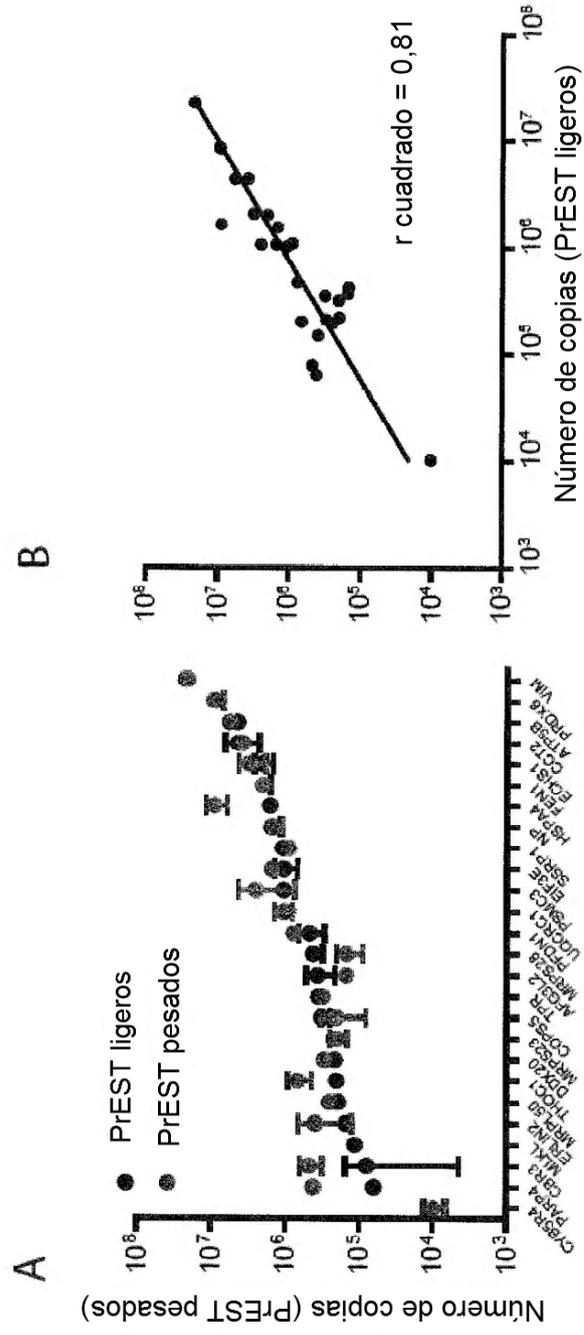


Figura 7



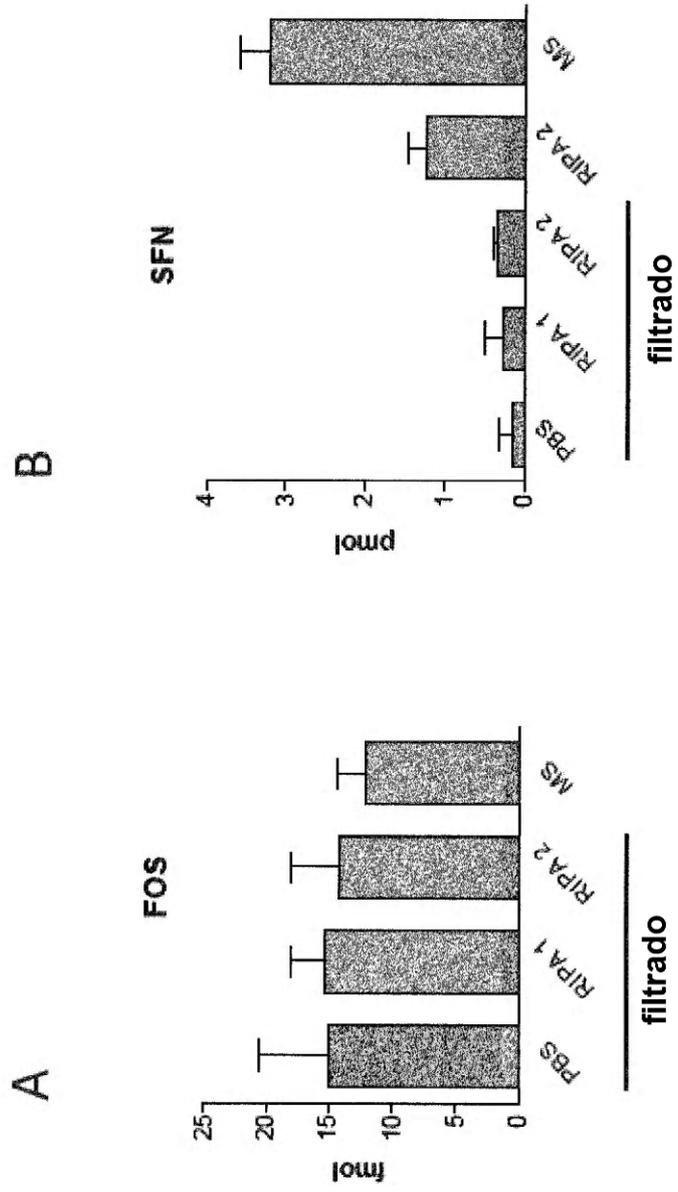


Figura 8

Figura 9

