

PROTEÓMICA Y MEDICINA

Fernando Vivanco. Fundación Jiménez Díaz, Madrid



Una de las consecuencias más beneficiosas derivadas del conocimiento del genoma humano ha sido el nacimiento y desarrollo de la proteómica. La descripción que hace el genoma de los organismos (Genómica) aporta un conocimiento esencial para el estudio de cualquier proceso biológico. Sin embargo, tal conocimiento es tan necesario como insuficiente, porque las funciones celulares dependen en última instancia de los productos finales de los genes: las proteínas y sus interacciones, asociaciones macromoleculares, vías de señalización, etc. De esta forma, la era post-genómica de la Medicina molecular se está desplazando rápidamente de los listados de genes, la transcriptómica y la genómica funcional, a la proteómica. Ésta ha dejado de ser (como en sus inicios) la producción de listas de proteínas que incrementan o decrecen su expresión como causa o consecuencia de diversas condiciones o alguna patología. Actualmente, la proteómica implica el conocimiento global de las proteínas, su estructura y funciones, interacciones y modificaciones post-traduccionales, rutas de señalización, localización intra e intercelular, etc., para generar el mapa celular que permita entender los procesos biológicos en condiciones normales y patológicas (lo que ha venido a denominarse la biología de sistemas). En suma, la proteómica es una actividad científica lo suficientemente madura como para ser aplicada a la clínica humana con notables beneficios.

El estudio de los proteomas, o de los perfiles proteicos de los líquidos biológicos (ver más adelante), presenta un grado de complejidad considerable que

no existe en los estudios genómicos, debido fundamentalmente a las siguientes causas:

a) El amplísimo rango dinámico de la concentración de proteínas en tejidos y especialmente en el plasma humano. Entre la concentración de albúmina (milimolar) y del TNF- α (fentomolar) hay una diferencia de 12 logs.

b) La presencia de diferentes formas moleculares de cada proteína. Muchos genes producen múltiples mRNAs, mediante diversos mecanismos (*splicing*, *mRNA editing*, modificaciones co- y post traduccionales). En promedio, en humanos, se estima que un gen genera al menos 6-8 proteínas, por lo que considerando ~35.000 genes, pueden esperarse un mínimo de 200-300.000 proteínas. Obviamente, no todas las proteínas estarán presentes en todas las células, tejidos, etc., pero el plasma contiene una gran representación de muchas de ellas.

c) Diversidad estructural. Dada su presencia en distintos entornos (solubles, membranas, matriz, etc.) las proteínas son de naturaleza muy diversa. Sólo en tamaño pueden variar desde 50 aminoácidos (hay ~2.500 anotados en el Swiss-Prot) hasta casi 10.000 aminoácidos, como la Nesprina 1 (8746 aa).

d) Una gran mayoría de las proteínas no actúan solas sino en asociación con otras, creando, por un lado, asociaciones macromoleculares y, por otro, participando en multitud de rutas de señalización, metabólicas, etc. El conocimiento de las interacciones proteicas y la descripción del mapa celular es esencial para entender la fisiopatología de cualquier sistema biológico.

Sin embargo, el conocimiento de los proteomas ofrece la visión óptima de lo que ocurre en la célula, tejido, líquido fisiológico....

PROTEÓMICA CLÍNICA

A pesar de la complejidad antes citada, el rápido y notable desarrollo de la proteómica en los últimos años ha conducido a su aplicación en los procesos patológicos o proteómica clínica. La proteómica está especialmente "capacitada" para el estudio de la enfermedad porque se puede aplicar a las células, tejidos y líquidos biológicos. La aplicación clínica de la proteómica implica el uso de las tecnologías proteómicas para todos aquellos aspectos que puedan beneficiar a los pacientes, desde la búsqueda de nuevos marcadores de enfermedad o de perfiles proteicos diagnósticos, hasta la descripción de factores predictivos

y preventivos que permitan un tratamiento y seguimiento terapéutico personalizado.

La proteómica clínica se ocupa de la identificación sistemática y exhaustiva, a gran escala, de patrones proteicos de enfermedad y de la aplicación de esos datos a los pacientes (estudio de la enfermedad, susceptibilidad, prevención, selección de terapias, seguimiento de los tratamientos, etc.). Dada la complejidad de los procesos patológicos, la integración de las nuevas tecnologías proteómicas con las bases de datos clínicas y las técnicas analíticas clásicas, sería altamente beneficioso para la comprensión y el tratamiento de la enfermedad. Entender, por ejemplo, las alteraciones tempranas en diversas enfermedades (típicamente en cáncer), produce una notable mejora en la prevención e identificación de pacientes con riesgo y/o en estado preclínico. La proteómica clínica pretende definir patrones de proteínas que puedan generar información de valor clínico acerca de la susceptibilidad, diagnóstico, pronóstico y terapia de la enfermedad. Testar hipótesis clínicas relevantes, frente a conjuntos de datos (masivos) derivados de medidas procedentes de técnicas de alto rendimiento, como las proteómicas, y poblaciones humanas fenotipadas correctamente, es una función esencial de la proteómica clínica.

Para que la proteómica clínica impacte de manera real en la mejora de la salud debe descubrir y seleccionar patrones proteicos, validarlos en estudios muy amplios y trasladarlos a la práctica clínica.

Aunque no se describe aquí, no debe olvidarse que el aprovechamiento óptimo de la relación entre patrones proteicos y enfermedad depende, a su vez, de la integración con la información sobre la variación genética y la expresión génica (*arrays* de DNA, genotipado, SNPs, etc.).

FINALIDAD DE LA PROTEÓMICA CLÍNICA

La finalidad de la proteómica clínica está, actualmente, centrada en al menos los siguientes tipos de actividades:

- 1) Búsqueda e identificación de biomarcadores individuales, siguiendo la metodología proteómica clásica de separación de proteínas (2-DE, cromatografía multidimensional y electroforesis capilar) e identificación por espectrometría de masas (MS; MALDI-TOF, ESI-Q-TOF o ESI-Trap, MALDI-TOF/TOF). Su estrategia fundamental es el análisis de la expresión diferencial de proteínas

entre controles y muestras patológicas, de proteomas completos (lisados celulares) o subproteomas específicos (mitocondrias, membranas, etc.). Un aspecto fundamental es la determinación de las modificaciones postraduccionales y su implicación en patología. En muchos casos, los biomarcadores se convierten simultáneamente en nuevas dianas terapéuticas.

2) Estudio sistemático, a gran escala, para la identificación de patrones o perfiles proteicos (huellas dactilares proteicas), con carácter diagnóstico (discriminatorio). Este estudio se puede llevar a cabo mediante tres estrategias:

a) En suero/plasma de los pacientes con la patología objeto de estudio (cardiovascular, neurodegenerativa, infecciosa, cáncer, etc.) y alternativamente en otros líquidos fisiológicos (lavado broncoalveolar, cefalorraquídeo, orina). Se utiliza la técnica de SELDI-TOF de alta resolución (SELDI-Q-Star MS/MS) para la generación de los perfiles proteicos, utilizando los biochips de Ciphergen. Debe considerarse que para un uso clínico óptimo, un biomarcador debe ser medible en suero (alternativamente en orina, o saliva). Los escasos biomarcadores identificados y aprobados (por la FDA) en los últimos años son un reflejo de la estrategia habitual de testar una proteína individual, análisis en pequeña escala, etc. Aunque existen excepciones (paraproteínas de los mielomas, déficit de $\alpha 1$ -antitripsina) es muy posible que no existan marcadores únicos de las enfermedades complejas, por lo que la estrategia proteómica de estudiar paneles proteicos supone un paso cualitativo de gran trascendencia.

b) En tejidos, mediante la obtención de perfiles o imágenes (mapas) bidimensionales de proteínas (MS-Imaging), aplicando directamente la MS sobre cortes histológicos del tejido (cortes congelados y/o teñidos), utilizando el MALDI-TOF o el MALDI-TOF/TOF. Esta tecnología puede utilizarse independientemente o en combinación con el escáner molecular, para el mapeo de las proteínas a través de sus péptidos. Un complemento esencial para el MS-Imaging es su aplicación sobre células o grupos de células procedentes del tejido, mediante microdissección por captura con laser (LCM). Además, esta técnica se aplica para localizar y mapear en el tejido moléculas pequeñas, particularmente fármacos, de gran interés para la industria farmacéutica.

c) La obtención de perfiles proteicos, tanto de suero (u otros líquidos fisiológicos) como de células o tejidos, puede también obtenerse a par-

tir de la utilización de chips (o micromatrices) de proteínas. Este es el sector de mayor crecimiento en el mercado proteómico, especialmente en el área de la proteómica funcional (análisis de las propiedades bioquímicas y celulares de las proteínas). No deben confundirse los chips de proteínas (micromatrices de proteínas capturadas o depositadas en una superficie) con la miniaturización (y automatización) de las distintas técnicas bioanalíticas proteómicas (2-DE, electroforesis capilar, cromatografías) que se están desarrollando en formato tipo chip (*"the lab-on-a-chip"*) basados en la tecnología de microfluídica. Los chips de proteínas pueden clasificarse según muchos parámetros (modo de fabricación, química de la superficie, especificidad, densidad, etc.). Sin embargo, es mejor sistematizarlos en dos grandes grupos, en base a lo que capturan o analizan:

Chips analíticos: Se utilizan para determinar las proteínas presentes en una muestra (perfil de expresión) y su cuantificación. La disolución de la mezcla de proteínas a analizar se aplica sobre el chip, que contiene agentes de captura, que actúan como sondas. Los agentes de captura más conocidos son los anticuerpos (Ab) en sus distintas modalidades: monoclonales completos, Fab, scFv, afibodies (Ab sintéticos), dominios de Ab, etc. El uso de Ab no está exento de problemas: inespecificidad, reacciones cruzadas, orientación, afinidad, etc. Otra alternativa son los chips de antígeno para analizar el perfil sérico de Ab (y autoAbs en enfermedades autoinmunes), como los chips de alérgenos para la detección de IgEs en respuestas alérgicas.

Recientemente se está generando una plétora de agentes de captura distintos a los Abs, como péptidos, moléculas orgánicas, polímeros sintéticos, oligonucleótidos, aptámeros, ribozimas, trinectinas (derivados de la fibronectina), MIPs (*molecular imprinted polymers*: compuestos que se polimerizan sobre las proteínas creando moldes para su uso posterior), etc. A pesar de esta diversidad, la captura de las proteínas (y los métodos de detección) constituye el gran problema de los chips proteicos: no existen todavía agentes de captura de alta calidad que puedan inmovilizarse en la superficie de un chip y que permitan unir, detectar y cuantificar una mezcla compleja de proteínas. Este problema está retrasando no sólo el desarrollo de los chips proteicos, sino de la industria proteómica en general. La investigación proteómica está generando multitud de biomarcadores potenciales y nuevas dianas terapéuticas, pero tiene el cuello de botella del *screening* y la validación de fármacos y los chips de proteínas pare-

cen el método adecuado para solucionarlo. Los chips de fase reversa son de especial relevancia en Medicina. Muestras de biopsias se lisan y se depositan en diluciones sucesivas como microgotas en el chip (utilizando los mismo robots para la fabricación de los chips de DNA), que son luego testados con Abs validados específicos de kinasas, factores de crecimiento, etc, obteniéndose información cualitativa y cuantitativa, para cada paciente.

Chips Funcionales: Se utilizan para determinar actividades bioquímicas e interacciones moleculares. Las proteínas de interés se inmovilizan en el chip y se analizan sus funciones biológicas e interacciones, incubándolas con distintos ligandos.

3) Utilización de herramientas bioinformáticas (y bioestadísticas) para el almacenamiento y procesamiento analítico de los datos masivos generados por proteómica. Su finalidad esencial es el desarrollo y aplicación de métodos de jerarquización y *clustering*, visualización de datos, algoritmos estadísticos multivariable, para definir los perfiles específicos de las distintas patologías. Una necesidad inmediata es la creación de bases de datos de proteínas implicadas en las distintas patologías, sus interacciones, participación en rutas de señalización, etc., así como su relación con los datos genómicos. La aplicación de estos métodos a gran escala no está desarrollada y debe implementarse para poder definir claramente cuál es el patrón "normal" (no existe actualmente) y poder discriminar, posteriormente, los patológicos y su gradación. Esto es esencial para poder definir y establecer test multiparámetros. Además, hay que desarrollar métodos que integren de forma efectiva la información procedente de diversos niveles biológicos (perfiles proteicos, niveles o concentraciones, imágenes histológicas, datos clínicos, etc.). Es evidente que el análisis masivo de datos y la integración de los distintos niveles de información, sólo puede llevarse a cabo mediante una plataforma bioinformática proteómica, que constituye uno de los elementos claves de cualquier unidad de proteómica clínica.

4) Creación de un Banco de Muestras Clínicas (suero, tejidos). Tanto los ensayos clínicos como los estudios de validación de biomarcadores o de perfiles necesitan colecciones amplias de muestras biológicas y sus protocolos de obtención y recogida de muestras, conservación, uso, etc. Tales estudios son particularmente necesarios para demostrar la especificidad diagnóstica del perfil proteico, esto es, su ausencia en los controles (incluyendo otras enfermedades) y su presencia en la enfermedad estudiada. Se contemplan dos tipos de grupos de muestras:

- a) Grupos reducidos de muestras clínicas muy bien caracterizadas, para la fase de descubrimiento de biomarcadores o perfiles y estudios piloto.
- b) Grandes colecciones de muestras clínicas bien caracterizadas, necesarias para la validación de los biomarcadores o perfiles y el progreso hacia su uso clínico de rutina.

CARÁCTER MULTIDISCIPLINAR DE LA PROTEÓMICA CLÍNICA

Con lo expresado hasta aquí, parece obvio que la proteómica clínica debe llevarse a cabo por equipos multidisciplinares, con clínicos, anatomopatólogos, químicos de proteínas, expertos en MS, bioinformáticos, farmacólogos, etc. En ocasiones, los análisis proteómicos proporcionan una avalancha de datos de difícil evaluación. Una de las funciones esenciales de los equipos multidisciplinares es la priorización de patrones candidatos que han de ser validados. La fiabilidad de los datos, avalados en las bases de datos, su potencial valor clínico y un riguroso análisis bioinformático, son algunos de los criterios que deben seguirse en la priorización.