

Fundamentos y Prospecciones del Paradigma de las Ciencias Ómicas en la Salud Humana (Segunda Parte)

Guadalupe Prado Flores¹ y Arturo César García Casillas²

Resumen. *La complejidad del flujo vital integra cualidades dinámicas, adaptativas, eficientes, jerárquicas y evolutivas, las cuales se pueden considerar con alto nivel de organización en las esferas cuántica, física y química. No sólo eso, hay una manifiesta emergencia hacia los niveles de la autonomía y la cognición que los organismos vivos expresan en su autopoiesis. En esa red de estructuras y funciones, el ADN constituye la plataforma fundamental de la información. En el campo actual de las ciencias biológicas progresa una corriente de las ciencias del genoma con la vinculación y aplicación de la bioinformática de vanguardia. Este sistema se integra con atributos multidisciplinario instrumental y fenomenológico. Con este conjunto se viene edificando un constructo teórico-metodológico-bioinformático relevante en el presente y con gran significado en la prospección del conocimiento de la vida planetaria. La biología de sistemas integra el conocimiento especializado de la genómica, epigenómica, transcriptómica, proteómica, metabolómica e interactómica como plataforma a otras disciplinas ómicas. Este*

¹ Profesor-Investigador, Departamento de Producción Agrícola y Animal, Universidad Autónoma Metropolitana, Unidad Xochimilco, e-mail: gprado@correo.xoc.uam.mx

² Investigador Posdoctoral, Maestría Interinstitucional en Producción Pecuaria, Universidad de Colima, e-mail: cesargarciasillas@hotmail.com

nuevo paradigma encauzado hacia la síntesis, está orientado fundamentalmente en la comprensión del dúo salud-patología, componente inexorable del binomio vida-muerte en el planeta.

Palabras Clave: *epigenómica, transcriptómica, proteómica, interactómica.*

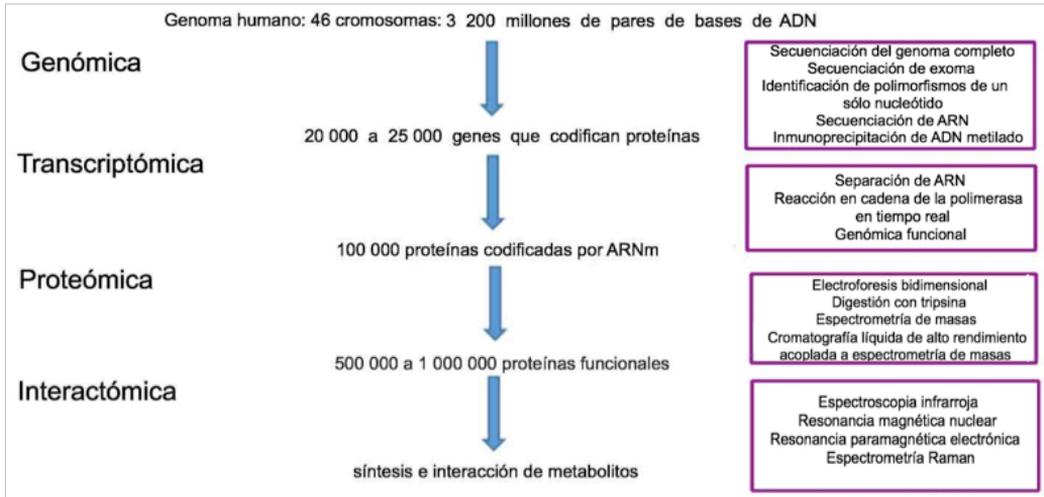
Abstract. *The complexity of the vital flow integrates dynamic, adaptive, efficient, hierarchical and evolutionary qualities that can be considered with high level of organization in the quantic, physical and chemical structures. Not only that; there is a manifest emergency towards the levels of autonomy and cognition, which living organisms express in their autopoiesis. The specie Homo sapiens has specific attributes; which them, the specie has the privilege of studying himself and the world. In the current field of biological sciences there is an integrative current of the genome sciences with the hand of the applicability from the bioinformatics vanguard. Contextually, it shows an instrumental and phenomenological multidisciplinary development. With this set, a cognitive-technological network has been built as a relevant theoretical-methodological-bioinformatics construct in the present and with a great significance in the prospection of the knowledge on the planet life. These biological systems integrate the specialized knowledge of genomics, epigenomics, transcriptomics, proteomics, metabolomics and interactomics as a platform for other omic disciplines. This new paradigm is directed towards synthesis, it is also oriented fundamentally in the understanding of the health-pathology, an inexorable component of life-death on the Earth.*

Keywords: *genomics, epigenomics, transcriptomics, proteomics, interactomics.*

INTRODUCCIÓN

A partir de los datos obtenidos del Proyecto del Genoma Humano (PGH), se han desarrollado herramientas de alta resolución como la secuenciación masiva (*International Human Genome Sequencing Consortium*, 2004). El estudio de 1,000 genomas en el año 2010 se pudo lograr gracias a tecnologías de alto rendimiento, como los secuenciadores de ADN que podían leer 250 mil millones de nucleótidos en una semana. Actualmente se puede lograr esta hazaña científica en un día (Heard, 2016). Después del estudio con el ADN, se han atendido procesos de involucramiento de diversos tipos de ARN; adicionalmente se ha hecho con proteínas y procesos como la metilación del ADN. Entre las técnicas y métodos científicos de nueva generación (Figura, 1) se pueden citar: i) detección de variantes; ii) secuenciación de genoma completo; iii) secuenciación del exoma; iv) identificación de polimorfismos de un solo nucleótido, SNPs; v) identificación de inserciones y deleciones; vi) identificación de inversiones; vii) identificación de variantes en el número de copias, CNVs, pérdida o ganancia de material genético; viii) secuenciación de ARN (ARN-seq); ix) inmunoprecipitación de cromatina; x) inmunoprecipitación de ADN metilado (MeDIP-seq), y xi) secuenciación de ADN tratado con bisulfito de sodio (WGBS) (Macaulay y Voet, 2014; Sims *et al.*, 2014).

Figura 1. Métodos científicos y técnicas relacionados con cuatro ciencias ómicas



Fuente: Modificado a partir de Plaza *et al.*, 2017.

Todos estos estudios con dichas herramientas han dado lugar al nacimiento de las “ciencias ómicas” (Ruiz *et al.*, 2014). Con esta denominación se evalúa al genoma, proteoma, epigenoma, transcriptoma, metaboloma e interactoma de manera integral (Plaza *et al.*, 2017). Su análisis se refiere a la caracterización y cuantificación de la totalidad de las moléculas biológicas que forman ciertas estructuras, que tienen determinadas funciones, las cuales están en integración constante en un organismo. Este paradigma puede considerarse como una visión compleja en la biología de sistemas.

Genómica

El término genoma alude a la constitución de todos los nucleótidos que conforman genes codificantes y no codificantes de proteínas, secuencias repetidas, regiones intergénicas que componen el ADN de los cromosomas nucleares y el único cromosoma mitocondrial (Paterson y Kolata, 2017). Contiene las claves de la herencia y su transmisión, incluidas aquellas características que confieren resistencia o susceptibilidad a cierto proceso patológico (Seehausen *et al.*, 2014).

La genómica se orienta a la evaluación de manera integral del ADN en un solo experimento, ya que estudia el genoma de un organismo o microorganismo apoyándose en herramientas de alto rendimiento y de última generación (Plaza *et al.*, 2017). Su estudio se ubica en las variantes genéticas, como los polimorfismos de un solo nucleótido, variantes en el número de copias, deleciones y ganancia de fragmentos cromosómicos (Seehausen *et al.*, 2014).

Sus herramientas fundamentales parten de la reacción en cadena de la polimerasa (PCR) en tiempo real, los microarreglos y la secuenciación masiva (Paterson y Kolata, 2017). Estas metodologías han evolucionado hasta la tercera generación, basadas en los conocimientos científicos más desarrollados, y bajo los parámetros de horizontalidad y profundidad; poseen recursos de la máxima precisión posible, así como de la vanguardia de las ciencias de la información (Macaulay y Voet, 2014). De manera sucinta se pueden citar: *Roche 454*; *Illumina*; *Applied Biosystems SOLID*; *Life Technologies Ion Torrent/Ion Proton*, y *Complete Genomics* (van Dijk *et al.*, 2014).

Entre las empresas-laboratorios que realizan secuenciación de tercera generación se pueden citar: i) *Pacific Biosciences* (2017), con tecnología SMRT para secuenciación de moléculas individuales en tiempo real. Combinan tecnología de semiconductores, fotónica y biotecnología con fluoróforos; ii) *GenapSys* (2017), se basan en la termosecuenciación; iii) *Genia's* (2017), con el empleo de nanoporos embebidos en una bicapa lipídica y electrodos que detectan cambios en la corriente eléctrica; iv)

Microarray Lab (2017) que poseen colecciones ordenadas de sondas inmovilizadas en una superficie sólida. Cada grupo de sondas hibrida con el problema a analizar; v) Affimetrix (2017) con el sistema *GeneChip* que utiliza un proceso de fotolitografía; vi) illumina (2017), quienes hacen microarreglos para genotipificación. Pueden analizar 12 muestras con una densidad de 90,001 a 250,000 esferas; vii) Agilent (2017) con microarreglos que involucran la síntesis *in situ* de oligonucleótidos de 60 bases mediante arreglos de hibridación comparativa del genoma.

Epigenómica

La epigenómica se encarga del conjunto de mecanismos que no afectan la estructura de los genes, pero que sí modifican su expresión (Hughes y Lambert, 2017). El conjunto de los factores epigenéticos que inciden en el estado de una célula se denomina epigenoma (Carrer y Wellen, 2015). Factores físicos como la temperatura, la duración de horas luz-oscuridad; químicos como la incorporación de fármacos; nutricionales como la dieta; perturbadores como el estrés y muchos otros, generan mecanismos epigenómicos como los siguientes: i) metilación de la citosina del ADN; ii) modificación de histonas; iii) posicionamiento de nucleosomas en el ADN; iv) cambios post-traslacionales mediados por micro ARN (miARN), y v) la noción de la impronta genética (Zoghbi y Beaudet, 2016). Todos estos mecanismos guardan relación con cambios en la estructura de la cromatina.

Mediante el estudio del epigenoma completo se analizan los mecanismos de una o varias células, tejido, órgano u organismo en una condición, estímulo, nutriente o patología en particular (Zoghbi y Beaudet, 2016). Desde 1935 se empezaron a analizar comportamientos celulares anómalos y sucesivamente se fue conformando una hipótesis epigenómica integrada de conceptos y respuestas con las siguientes connotaciones: i) la expresión del ADN se modifica por factores internos y externos; ii) estos fenómenos se presentan en diferentes estadios del desarrollo; iii)

en órganos en desarrollo permanecen células embrionarias con características desdiferenciadas; iv) la actividad normal celular se desprograma y genera patrones neoplásicos del metabolismo; v) la expresión epigenética da una programación anormal de los genes durante la diferenciación; y vi) adicional a los cambios fenotípicos, hay cambios genotípicos (Carrer y Wellen, 2015; Kinnaird *et al.*, 2016; Korbel y Roberts, 2017).

Los mecanismos epigenómicos se estudian a partir de la purificación del ADN, lo cual se hace por electroforesis capilar y/o HPLC. El contenido de metilaciones de citosinas en las islas CpG del ADN se estiman mediante inmunoprecipitación de un solo gen (*Single-gene-chromatin immunoprecipitation*) con antígenos específicos para factores de transcripción o proteínas. También se emplean las técnicas de RLGs (*Restriction Landmark Genome Scanning*), tratamientos con bisulfito de sodio para determinar las transformaciones de las citosinas a uracilo. Se analizan los patrones de metilación por secuenciación del genoma y las amplificaciones se hacen por PCR. Las modificaciones en las histonas se analizan por mapeo (*DNA Methylomes and Histones Modification Maps*) y Espectrometría de Masas (MS). Las modificaciones en zonas no codificantes de ARN se estudian mediante la extracción de ARN, seguida de microarreglos. Los polimorfismos de un solo nucleótido se identifican por secuenciación polimórfica conformacional de cadena única o Cromatografía Líquida Desnaturalizante de Alto Rendimiento (DHPLQ). La Amplificación de Sitios Intermetilados se estima por AIMS y por la Hibridación Diferencial de Metilación (DMH). La actividad catalítica de las metiltransferasas y desmetiltransferasas se realiza por cinética enzimática (Esteller, 2007).

Se han encontrado abundantes respuestas epigenómicas en enfermedades multifactoriales; un ejemplo relevante lo constituye el cáncer mamario, en el cual hay numerosas distorsiones epigenéticas. Cuando se pone la atención solamente en la estructura y función del receptor estrogénico a ($ER\alpha$), así como su expresión, se aprecia que esta molécula manifiesta respuestas que se valoran como mecanismos epigenómicos. El

ER α participa en la elongación transcripcional como factor de transcripción (FT). Su acción la realiza al interactuar con el complejo del Factor-B mediante la fosforilación de la Ser 2 en el dominio del carboxilo terminal de la ARN pol II (Bedi *et al.*, 2015). Diversos autores han dado evidencia de la red de respuestas que este receptor genera en el ámbito epigenómico. Las más conspicuas son las siguientes: metilaciones específicas en el ADN, modificaciones en genes, modificación en patrones de acetilación y metilación de histonas, desprogramación de la cromatina, cambios en fragmentos no codificantes del ARN y presencia de SNP (Byler *et al.*, 2014). El gen de ER α tiene como genes diana a GREB 1 y pS2 (Feng *et al.*, 2014). Se asocia con el reclutamiento de varias enzimas modificadoras de histonas y remodelación de la cromatina. Las interacciones entre una chaperona de histonas, la SUPT6H, y la histona ubiquitinizada, H2Bub1, están asociadas a la desdiferenciación y malignidad del tumor (Bedi *et al.*, 2015). En el cáncer mamario se localizan rasgos específicos en la metilación del ADN: hipermetilación en supresores de tumores, dominios parcialmente metilados en oncogenes y/o epitelio luminal muy metilado (Holm *et al.*, 2016). También se han evidenciado factores de transcripción ausentes o modificados (Bedi *et al.*, 2015); sobreexpresión de receptores estrógenicos y, en otros casos, ausencia de ER α , RP y HER2 (Bodenstine *et al.*, 2016); mutaciones múltiples, SNPs, entre ellos la Thr 394 en células de sangre periférica (Li *et al.*, 2012), o los polimorfismos asociados a la densidad mamaria (Yong *et al.*, 2010).

Las alteraciones epigenéticas contribuyen a la tumorigénesis. Vogelstein *et al.* (2013) han encontrado 140 genes asociados en esta megadisfunción. Sus datos apuntan en mutaciones en regiones no codificantes de naturaleza *cis* que tienen relación *trans*, es decir, secuencias de nucleótidos que repercuten en proteínas; transformaciones en histonas con lisinas específicas metiladas y mutaciones que rompen el equilibrio entre metiltransferasas y desmetiltransferasas (Zoghbi y Beaudet, 2016).

El estudio de los mecanismos epigenómicos ha llevado tanto a los ensayos multigenes como a terapias de resultados muy satisfactorios

(Gyorffy *et al.*, 2015; Naoi y Noguchi, 2016). Hoy en día las denominadas “plataformas genómicas” abordan promisoriamente el problema: i) *OncoType* estudia 21 genes con *Quantitative real-time RT-PCR*. Validan sus parámetros de recurrencia con modelos de regresión (Flanagan *et al.*, 2008); ii) *ProsignaPAM30* utiliza *Quantitative real-time RT-PCR* y microarreglos en 50 transcritos con *NanoString's Technologies*, que es una plataforma transcripcional (Wallden *et al.*, 2015), y iii) *MammaPrint* analiza 70 genes mediante exploraciones transcriptómicas y propone una estratificación de tratamientos (Cusumano *et al.*, 2014). Este campo de estudio es un ejemplo de acciones epigenómicas cuyos nodos inciden en la genómica y generan una red interactiva de respuestas sobre el transcriptoma, manifiesto en el proteoma y, a su vez, en el metaboloma (Gyorffy *et al.*, 2015).

Transcriptómica

A partir de la estructura primaria del ADN se genera el ARN, proceso que tiene el mismo lenguaje químico, a excepción de la timina del ADN, que cambia por uracilo en el ARN. Es un proceso central que permite comprender la manera en que se relaciona el genotipo con el fenotipo, ya que esta transferencia de la información ofrece gran plasticidad (Bartlett y Stirling, 2003). Refleja el hecho de que todas las células contienen el mismo ADN, pero que se diferencian y dan lugar a órganos con tejidos y funciones diferentes, específicos e integrados (Zhegunov, 2012).

Edmonds *et al.* (1971) describieron la presencia de secuencias de homopolímeros de adenina (poli-A), unidos de forma covalente en el extremo 3'; ahora se sabe que son agregados al final de la transcripción, y que controlan la estabilidad de los propios transcritos, además participan en la exportación del ARNm del núcleo al citoplasma y tienen influencia en la traducción en el ribosoma. Por el año 1977, Rungger y Crippa descubrieron que los transcritos primarios de eucariontes debían pasar por un proceso de maduración que implica la remoción de intrones

y conservación de exones mediante el mecanismo de corte y empalme (Rungger y Crippa, 1977).

La transcripción del ADN a ARN se descifra en dirección 5' a 3', y puede leerse en agrupaciones de tres bases llamadas codones, que determinan la secuencia de los aminoácidos (aa) de la proteína en la dirección amino terminal a carboxilo terminal (Wray *et al.*, 2003). Al comprender que el ADN, por medio de sus transcritos de ARN, codifica para proteínas, se acepta que todas las células de un organismo tienen el mismo genoma, pero no tienen el mismo proteoma. Por tanto, el proceso de transcripción, modulado por factores internos y externos, le confiere plasticidad y constituye el vínculo entre el genotipo y el fenotipo (Rio, 2015). Al conjunto de los ARNm expresados se le conoce como transcriptoma, y de manera indirecta refleja al proteoma, lo cual no es estricto, ya que es importante reconocer que hay zonas de regulación previas a la traducción (Gupta y Warner, 2014). Adicionalmente, pueden darse mutaciones, SNP o cambios en su nivel de expresión (Sandelin *et al.*, 2007).

La transcriptómica tiene una relevancia específica, pues la generación de los ARNm que codifican para proteínas, así como los ARNnc que regulan respuestas tan variadas como numerosas, tienen un lugar irremplazable en el funcionamiento básico de los organismos (Hughes y Lambert, 2017). Las mutaciones o los SNP pueden modificar estos mecanismos esenciales con detrimento de la normalidad y están implicados en procesos degenerativos y patologías multifactoriales (Bass, 2015).

En la actualidad existen diversas herramientas moleculares para evaluar el transcriptoma. Las más importantes pueden considerarse a las siguientes metodologías: i) *Northern Blot*, evalúa los cambios en la expresión mediante la transferencia del ARN resuelto en geles de agarosa por electroforesis, en un medio desnaturante a un soporte sólido, seguido de una hibridación con sondas de ADN marcadas con ^{32}P , lavados y revelados por autorradiografía (Pall y Hamilton, 2008); ii) *Quantitative real-time RT-PCR*. Mediante transcriptasa reversa y oligonucleótidos (oligo dT) o cebadores específicos es posible sintetizar las cadenas

complementarias de ADNc, y después amplificar los genes de interés por PCR. El rendimiento debe ser evaluado en la fase exponencial de la reacción de la polimerasa (Bustin *et al.*, 2005); iii) *Fluidigm*. Consiste en generar matrices de reacción en escala de nanovolúmenes, y el flujo de los fluidos se controla con válvulas microscópicas. Pueden utilizar chips y pozos para 48 conjuntos de ensayos que produce matrices de 2,304 reacciones o 96 pozos que generan 9,216 reacciones (Jang *et al.*, 2011). De reciente manufactura existe el equipo C1™ *Single-Cell Auto Prep System* que puede tener aplicaciones en la expresión génica de hasta 96 genes, o 96 miARN o secuenciación de ARNm o ADN mediante NGS (Fluidigm, 2017); iv) *Functional genomics*. Es el estudio de la función de los genes a través de mediciones paralelas de expresión de sus genomas. Sus técnicas son el análisis serial de la expresión génica (SAGE), la secuenciación del ARN y los microarreglos; se basa en que 2.5% del genoma tiene secuencias conservadas que no codifican proteínas, pero codifican elementos funcionales, algunos son transcritos y otros tienen función reguladora (Morgan *et al.*, 2013). Con base en estos datos se generó el proyecto ENCODE (*ENCyclopedia Of DNA Elements*) con el fin de anotar el genoma humano con la función de los elementos que lo componen. Por este medio se conoce que 99% del genoma se encuentra dentro de 1.7 kbps de alguno de los episodios bioquímicos medidos por ENCODE (Consortium, 2004); v) *RNA-seq* incluye la secuenciación del ARN, *CAGE* (*cap analysis gene expression*) que identifica y cuantifica los extremos 5' de los ARN con capuchón, y *ARN-PET* (*RNA Paired-End tags*) que captura fragmentos producidos por una endonucleasa, la cual genera cortes frecuentes (Wang *et al.*, 2009), y vi) *Expression microarrays*. Detecta regiones de transcripción mediante microarreglos con diseño de mosaico. Con esa estrategia han localizado gran cantidad de transcritos de función desconocida (Consortium, 2015). Un modelo más reciente, el *GeneChip™ Human Transcriptome Array 2.0*, estudia la expresión diferencial de isoformas generadas por *splicing* alternativo (Thermo Fisher Scientific, 2017). Estos fenómenos están relacionados con los ARNm y ARNc inducidos

por algún estímulo o procesos patológicos, y las herramientas metodológicas descritas pueden conducir a la identificación de marcadores y potenciales blancos terapéuticos.

Proteómica

A la proteómica, junto con la transcriptómica, se les llama el genoma dinámico, ya que cada una, con sus plataformas tecnológicas, evalúa la manera como el genoma responde a un estímulo o patología en particular, o cuáles son sus expresiones basales en una condición dada (Wang *et al.*, 2009; Marx, 2014). Se habla de la proteómica cuando se analizan las proteínas traducidas en un solo genoma (Catherman *et al.*, 2014).

La proteómica ofrece altas expectativas en el entendimiento de las bases moleculares de las patologías al proponer nuevos marcadores para el pronóstico, diagnóstico o respuesta a tratamientos, así como nuevos blancos terapéuticos que pueden usarse en el diseño de fármacos (Gregorich y Ge, 2014). En concreto, la proteómica clásica se orienta en la búsqueda de marcadores; a su vez la proteómica estructural se refiere al diseño de fármacos (Catherman *et al.*, 2014).

En la búsqueda de marcadores se siguen protocolos del procesamiento de la muestra para separar proteínas específicas mediante la secuencia en las etapas de separación, identificación y validación (Gregorich y Ge, 2014). La extracción es el paso inicial de la separación, y depende del tipo de tejido al que se trata con un amortiguador idóneo. La lisis celular expone todos los componentes celulares y deben retirarse sin afectar la solubilidad de la proteína de estudio. Es muy frecuente utilizar detergentes, agentes caotrópicos, reductores o alquilantes que ayudan en la solubilización de las proteínas de interés, y la elección ha de realizarse bajo un cuidadoso marco de conocimiento de las interacciones que no perturben el objetivo del proceso (Stone *et al.*, 2017).

Se deben tomar en cuenta las condiciones de temperatura, pH, salinidad, iones metálicos y cofactores en cada etapa de separación, la cual se hace generalmente mediante centrifugación diferencial. Después de la separación se procede a la purificación que generalmente se realiza en gradientes de sacarosa o de cloruro de cesio (Orchard *et al.*, 2003). La cuantificación de la proteína total es un paso de análisis espectrofotométrico. Las técnicas clásicas de separación de proteínas son la electroforesis unidimensional o bidimensional, la cromatografía líquida de alta resolución (HPLC), y entre las técnicas de mayor capacidad de resolución e identificación está la espectrometría de masas, el mapeo de su estructura primaria y la difracción por rayos X (Catherman *et al.*, 2014).

En la validación de biomarcadores se aplica un *Western Blot* que transfiere las proteínas separadas de un medio gelificado a una membrana de nitrocelulosa (Orchard *et al.*, 2003). Una proteína específica puede ser identificada en un *Western Blot* por inmunodetección. Para esto, se emplea un anticuerpo que localiza la proteína de interés y un segundo anticuerpo contra la especie en la que se produjo el primario a fin de localizar los complejos anticuerpo-antígeno. Ese anticuerpo secundario se puede marcar sobre una película radiográfica o una mancha de algún color en una membrana (Hirano, 2012).

La inmunohistoquímica es otra técnica con la cual se valida un biomarcador. Consiste en la relación estricta y específica de un anticuerpo específico con un antígeno específico. La reacción sólo es visible si el anticuerpo está acoplado a una sustancia que sea capaz de absorber o emitir luz o producir algún color (Marx, 2014). La técnica recurrente en estos campos de la validación es la inmunofluorescencia que puede cuantificarse (Laurinavicius *et al.*, 2016).

Es muy importante que el diseño de análisis de la proteína responda al objetivo inicial, además que la pulcritud de esta manualidad sea con controles positivos y/o negativos, y que en todo el proceso de separación, identificación y validación del biomarcador proteico se cubran las características analíticas de sensibilidad, especificidad, exac-

titud y precisión (Orchard *et al.*, 2003). De esta manera, la proteómica permite obtener firmas moleculares o la caracterización de proteínas y su interacción al modelar las redes de comunicación celular que explican los diferentes estadios celulares o la comparación entre casos y controles (Marx, 2014).

Interactómica

Las tecnologías ómicas han generado millones de datos que se evalúan mediante sistemas bioinformáticos y ofrecen respuestas de las alteraciones moleculares de innumerables patologías (Gregorich y Ge, 2014). Su orientación se ubica en las interacciones proteína-proteína, proteína-ADN, proteína-ARN, proteína-metabolitos, ADN-ARN y otras más (Catherman *et al.*, 2014).

Desde la visión de interactomas separados, la interactómica estudia el conjunto de interacciones entre distintas biomoléculas en un entorno determinado, centrado principalmente en las interacciones de proteínas (Wang y Qian, 2014). Hace la integración de estas redes en una matriz que estudia de manera conjunta un organismo, órgano, tejido o célula (Hakhverdyan *et al.*, 2015). Las estimaciones actuales sugieren que el interactoma humano está formado aproximadamente de entre 130,000 a 650,000 interacciones proteicas, de las cuales sólo un subconjunto ha sido identificado de manera experimental. Cada vez cobran más importancia las redes de genes, de genes con ARNs, de ARNnc con ARN de interferencia (ARNsi), entre otras (Stumpf *et al.*, 2008).

Los desafíos más acuciantes de la interactómica son el entendimiento de las patologías como resultado tanto de las perturbaciones ambientales, genéticas, epigenéticas, transcripcionales, como del impacto de la variabilidad genética de cada individuo en el contexto de sus interacciones (Kiemer y Cesareni, 2007). De esta manera, la biología de sistemas ayuda a comprender los fenómenos biológicos y se orienta

a una medicina preventiva, predictiva, personalizada y participativa (Hakhverdyan *et al.*, 2015).

Los métodos y técnicas experimentales para descubrir interacciones proteína-proteína son: i) *Yeast two-hybrid* (Y2H), que detecta interacciones entre proteínas X y Y, donde X está unida al dominio BD (promotor de la secuencia), el cual se une a la secuencia promotora (Bruckner *et al.*, 2009); ii) *Mass spectrometry*, que identifica la secuencia de miles de polipéptidos simultáneamente (Heck, 2008); iii) *Tandem Affinity Purification* (TAP) purifica complejos proteicos y remueve las moléculas contaminantes (Davis *et al.*, 2006); iv) la letalidad sintética que describe la interacción genética cuando una combinación de mutaciones entre dos o más genes o proteínas conduce a la muerte celular (Bruckner *et al.*, 2009), y v) los microarreglos de proteínas que dan cuenta de las proteínas que interactúan con un anticuerpo específico, ya que emiten fluorescencia e identifica cuántas y cuáles proteínas interactúan con el anticuerpo (Sun *et al.*, 2008).

Un interactoma se compone de nodos, enlaces y módulos (Hakhverdyan *et al.*, 2015). Los nodos pueden representar proteínas, genes, metabolitos, ARNs e incluso patologías y fenotipos. Dichos nodos, según su localización en el mapa topológico, pueden ser centrales, periféricos o cuellos de botella (Handley y Behler, 2014). Los enlaces representan las interacciones de los nodos por contactos físicos proteína-proteína, o relaciones de co-expresión de genes o acoplamiento metabólico (Kiemer y Cesareni, 2007). También pueden representar patologías basadas en un origen genético común o que comparten características relevantes (Bruckner *et al.*, 2009). Los módulos son las áreas de nodos más densas dentro del interactoma, es decir, son subconjuntos de nodos que tienen más interacciones entre sí que con el resto de los componentes. Se pueden clasificar en topológicos, funcionales y patogénicos (Aitchison y Rout, 2015).

Las bases de datos de interacciones se pueden clasificar en: i) las que contienen interacciones comprobadas experimentalmente, p. ej. *Bio-*

molecular Interaction Network Database (BIND, 2017), *Database of Interacting Proteins* (DIP, 2017), y *The Molecular INTeraction database* (MINT, 2017); ii) las que presentan interacciones deducidas por métodos de predicción computacional, p. ej. *Domain Interaction MAp* (DIMA, 2017), *Interacting Protein Domains* (InterDom, 2017), *Structural Classification of Protein–Protein Interfaces* (SCOPPI, 2017), y iii) las que almacenan e integran los dos tipos anteriores, p. ej. *Search Tool for the Retrieval of Interacting Genes/Proteins* (STRING, 2017).

Las redes se construyen bajo dos conceptos: el análisis topológico que cumpla los conceptos de la teoría de redes y la interpretación biológica que pueda representar algún proceso funcional (Kiemer y Cesareni, 2007). A su vez, se mide la significancia estadística del modelo aplicando tanto el modelo, global como el local. La conectividad global (CG) consiste en la suma de los nodos y los enlaces que componen el interactoma. Mediante el cálculo de una P empírica menor a 0.05, el interactoma es significativo respecto al azar con una confianza de 95% (Zhang y Trinkle, 2015). Ya comprobada la significatividad estadística con listas de por lo menos 1,000 elementos, se determina mediante herramientas bioinformáticas su correspondencia a un proceso celular (Yamamoto *et al.*, 2009).

Las proteínas que componen cada módulo del interactoma generalmente se analizan por separado (Kiemer y Cesareni, 2007). Para este análisis se extraen las anotaciones de diferentes bases de datos, p. ej., *KEGG pathways* (Wrzodek *et al.*, 2013), *Gene Ontology (GO) Consortium* (Gene Ontology *et al.*, 2013), *ArrayXpath* visualiza datos de expresión génica provenientes de microarreglos de diferentes procesos biológicos (Salleh *et al.*, 2013), *COSMIC Cancer Gene Census* (Sondka *et al.*, 2017), entre otros.

La importancia de los interactomas es que se acercan con suficiente fidelidad a la muy compleja red estructural y funcional de los organismos en su totalidad (Hakhverdyan *et al.*, 2015). Entre las aplicaciones más importantes se cuenta con hallazgos de genes candidato de patologías como cáncer; también se dan evidencias de que las proteínas de

nodos centrales tienden a estar codificadas por genes esenciales en la fisiología celular (Kiemer y Cesareni, 2007). Al respecto, Xu *et al.* (2016) proponen un método de ponderación de red estructural para utilizar la información del complejo micro ARN-TF-ARNm en la identificación de subtipos de cáncer.

Con este modelo Xu *et al.* (2016) han estudiado el glioblastoma multiforme, un cáncer muy común y agresivo, y el carcinoma invasivo mamario. Las interacciones se recuperaron de diversas bases de datos de interactómicas al detectar proteínas humanas codificadas por genes *drivers* o candidatos de cáncer, las cuales tienen mayor cantidad de interacciones que las proteínas que no han sido caracterizadas como mutadas. Verificaron que las mutaciones en genes candidato para cáncer causan un fuerte impacto en la estructura de proteínas, y dicha información se ha asociado con la pérdida de la estructura tridimensional de las proteínas como un mecanismo molecular que está altamente relacionado con el fenotipo maligno.

Estos modelos integradores son una opción de gran trascendencia en la comprensión de patologías multifactoriales porque su planteamiento es sólido y amplio, y porque a partir de las relaciones encontradas se ofrecen propuestas bastante definidas en los tratamientos a seguir para mejorar la calidad de vida o la completa solución del problema de salud analizado.

Con las bases señaladas de la genómica, epigenómica, proteómica e interactómica se desarrollan disciplinas que abordan variados modelos sistémicos. Entre las de mayor avance se pueden citar las siguientes: metabolómica, medicina genómica, farmacogenómica, nutrigenómica, metagenómica, exposómica, secretómica, fenómica, citogenómica, citocinómica y lipidómica (Ramírez Bello, 2016). Cada una de estas disciplinas se conecta con otras con la generación de un interactoma mayor, de modo que la ciencia biológica, desde una óptica objetiva, puede validarse como un sistema complejo de conocimiento.

CONCLUSIONES

El campo de acción de las ciencias ómicas corresponde principalmente a procesos patológicos humanos, aunque también se aprecia su vinculación en patologías animales, y así se registra la mastitiómica. Asimismo, este paradigma es un recurso integrador en la promoción de la agricultura sustentable; de modo particular, la productividad alimentaria se ha orientado hacia el maíz. También se comienza a explorar sus beneficios en otras áreas como el deporte de alto rendimiento.

Se ratifica que su nivel científico avanza a una velocidad vertiginosa gracias a la conjunción de las ciencias de la información. Mediante sus conceptos, lenguajes y recursos se ha acelerado el desarrollo de las ciencias de la salud. Actualmente, el paradigma de las ciencias ómicas es una clara evidencia.

El proceso de hacer ciencia en el área biológica ha cubierto etapas secuenciales: del enfoque descriptivo pasó a un nivel de análisis para lograr ser interpretativo; más tarde, llegó a concebir explicaciones en una multitud de fenómenos. En esta apertura conceptual, metodológica y con la participación de las ciencias físicoquímicas, la ciencia biológica llegó a ser explicativa y fenomenológica. Con la integración de los matices anteriores, el momento actual manifiesta un auge aplicativo. De una manera evidente, las ciencias biológicas son generadoras de conocimiento. El siguiente paso puede ser tender puentes de reflexión y síntesis conjunta con las ciencias humanísticas en un proceso de complejidad creciente, de mayor amplitud, de inclusión ordenada en la construcción de transdisciplinabilidad.

BIBLIOGRAFÍA

- Affymetrix, 2017, "GeneChip Scanner 3000", en <https://www.affymetrix.com/>, consultado 01/08/17.
- Agilent, 2017, "Comparative genomic hybridization using oligonucleotide microarrays", en <http://www.agilent.com/labs/>, consultado 01/08/17.
- Aitchison, D. y M. Rout, 2015, "The interactome challenge", en *J Cell Biol* 211(4): 729-732.
- Bartlett, M. y D. Stirling, 2003, "A short history of the polymerase chain reaction", en *Methods Mol Biol* 226(1): 3-6.
- Bass, L., 2015, "Twenty years: a very short sequence in the RNA world", en *RNA* 21(4): 490-491.
- Bedi, U. *et al.*, 2015, "SUPT6H controls estrogen receptor activity and cellular differentiation by multiple epigenomic mechanisms", en *Oncogene* 34(4): 465-473.
- BIND, 2017, "Information about interactions between two biological 'objects', A and B, which could be protein, RNA, DNA, molecular complex, small molecule, photon (light) or gene", en <http://bind.ca/>, consultado 01/08/17.
- Bodenstine, M. *et al.*, 2016, "Nodal expression in triple-negative breast cancer: Cellular effects of its inhibition following doxorubicin treatment", en *Cell Cycle* 15(9): 1295-1302.
- Bruckner, A. *et al.*, 2009, "Yeast two-hybrid, a powerful tool for systems biology", en *Int J Mol Sci* 10(6): 2763-2788.
- Bustin, S. *et al.*, 2005, "Quantitative real-time RT-PCR a perspective", en *J Mol Endocrinol* 34(3): 597-601.
- Byler, S. *et al.*, 2014, "Genetic and epigenetic aspects of breast cancer progression and therapy", en *Anticancer Res* 34(3): 1071-1077.
- Carrer, A. y K. Wellen, 2015, "Metabolism and epigenetics: a link cancer cells exploit", en *Curr Opin Biotechnol* 34(1): 23-29.

- Catherman, D. *et al.*, 2014, "Top Down proteomics: facts and perspectives", en *Biochem Biophys Res Commun* 445(4): 683-693.
- Consortium, P., 2004, "The ENCODE (ENCyclopedia Of DNA Elements) Project", en *Science* 306(5696): 636-640.
- Consortium, T., 2015, "Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans", en *Science* 348(6235): 648-660.
- Cusumano, G. *et al.*, 2014, "European inter-institutional impact study of MammaPrint", en *Breast* 23(4): 423-428.
- Davis, P. *et al.*, 2006, "Protein complex compositions predicted by structural similarity", en *Nucleic Acids Res* 34(10): 2943-2952.
- DIMA, 2017, "Comprehensive resource for functional and physical interactions among conserved protein-domains", en <http://webclu.bio.wzw.tum.de/dima>, consultado 01/08/17.
- DIP, 2017, "Database catalogs experimentally determined interactions between proteins", en <http://dip.doe-mbi.ucla.edu/dip/Main.cgi>, consultado 01/08/17.
- Edmonds, M. *et al.*, 1971, "Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship", en *Proc Natl Acad Sci USA* 68(6): 1336-1340.
- Esteller, M., 2007, "Cancer epigenomics: DNA methylomes and histone-modification maps", en *Nature Publishing Group* 8: 286-299, en www.nature.com/reviews/genetics.
- Feng, Q. *et al.*, 2014, "An epigenomic approach to therapy for tamoxifen-resistant breast cancer", en *Cell Res* 24(7): 809-819.
- Flanagan, B. *et al.*, 2008, "Histopathologic variables predict oncoTYPE DX (TM) recurrence score", en *Mod Pathol* 21(10): 1255-1261.
- Fluidigm, 2017, "C1™ Single-Cell Auto Prep System", en <https://www.fluidigm.com/>, consultado 01/08/17.
- GenapSys, 2017, "The term "sequencing", en <http://www.genapsys.com/>, consultado 01/08/17.

- Gene, C. *et al.*, 2013, "Gene Ontology annotations and resources", en *Nucleic Acids Res* 41(Database issue): D530-D535.
- Genia's, 2017, "Integrated circuits in NanoTag chemistry for massively parallel single-molecule DNA sequencing", en <http://www.geniachip.com/>, consultado 01/08/17.
- Gregorich, R. y Y. Ge, 2014, "Top-down proteomics in health and disease: challenges and opportunities", en *Proteomics* 14(10): 1195-1210.
- Gupta, V. y J. Warner, 2014, "Ribosome-omics of the human ribosome", en *RNA* 20(7): 1004-1013.
- Gyorffy, B. *et al.*, 2015, "Multigene prognostic tests in breast cancer: past, present, future", en *Breast Cancer Res* 17(1): 11-16.
- Hakhverdyan, Z. *et al.*, 2015, "Rapid, optimized interactomic screening", en *Nat Methods* 12(6): 553-560.
- Handley, M. y J. Behler, 2014, "Next generation interatomic potentials for condensed systems", en *Eur Phys J B* 87(7): 152-162.
- Heard, E., 2016, "Epigènétique et cancer. Une brève histoire du cancer: gènétique et epigènétique", en <http://www.college-de-france.fr/site/edith-heard/course-2015-2016.htm>, consultado el 01/08/17.
- Heck, J., 2008, "Native mass spectrometry: a bridge between interactomics and structural biology", en *Nat Methods* 5(11): 927-933.
- Hirano, S., 2012, "Western blot analysis", en *Methods Mol Biol* 926(1): 87-97.
- Holm, K. *et al.*, 2016, "An integrated genomics analysis of epigenetic subtypes in human breast tumors links DNA methylation patterns to chromatin states in normal mammary cells", en *Breast Cancer Res* 18(1): 27-32.
- Hughes, R. y S. Lambert, 2017, "Transcription factors read epigenetics", en *Science* 356(6337): 489-490.
- illumina, 2017, "Microarray Solutions", en <https://www.illumina.com/>, consultado 01/08/17.
- InterDom, 2017, "Database of putative interacting protein domains derived from multiple sources, ranging from domain fusions (Rosetta Stone), protein interactions (DIP and BIND), protein complexes (PDB),

- to scientific literature (MEDLINE)", en <https://www.hsls.pitt.edu/obrc/index.php?page=URL1100033411>, consultado 01/08/17.
- International Human Genome Sequencing Consortium, 2004, "Finishing the euchromatic sequence of the human genome", en *Nature* 431(7011): 931-945.
- Jang, S. *et al.*, 2011, "Quantitative miRNA expression analysis using fluidigm microfluidics dynamic arrays", en *BMC Genomics* 12(1): 144-151.
- Kiemer, L. y G. Cesareni, 2007, "Comparative interactomics: comparing apples and pears?", en *Trends Biotechnol* 25(10): 448-454.
- Kinnaird, A. *et al.*, 2016, "Metabolic control of epigenetics in cancer", en *Nat Rev Cancer* 16(11): 694-707.
- Korbel, J. y C. Roberts, 2017, "A Convergence of Genetics and Epigenetics in Cancer", en *Cell* 168(4): 561-563.
- Laurinavicius, A. *et al.*, 2016, "Comprehensive Immunohistochemistry: Digital, Analytical and Integrated", en *Pathobiology* 83(2-3): 156-163.
- Li, L. *et al.*, 2012, "DNA methylation in peripheral blood: a potential biomarker for cancer molecular epidemiology", en *J Epidemiol* 22(5): 384-394.
- Macaulay, C. y T. Voet, 2014, "Single cell genomics: advances and future perspectives", en *PLoS Genet* 10(e): e1004126-e1004136.
- Marx, V., 2014, "Proteomics: An atlas of expression", en *Nature* 509(7502): 645-649.
- Microarray Lab, 2017, "Microarrays of immobilized oligonucleotide", en <http://brainarray.mbni.med.umich.edu/Brainarray/default.asp>, consultado 01/08/17.
- MINT, 2017, "Experimentally verified protein-protein interactions", en <http://mint.bio.uniroma2.it/>, consultado 01/08/17.
- Morgan, C. *et al.*, 2013, "Biodiversity and functional genomics in the human microbiome", en *Trends Genet* 29(1): 51-58.
- Naoui, Y. y S. Noguchi, 2016, "Multi-gene classifiers for prediction of recurrence in breast cancer patients", en *Breast Cancer* 23(1): 12-18.

- Orchard, S. *et al.*, 2003, "The proteomics standards initiative", en *Proteomics* 3(7): 1374-1376.
- Pacific Biosciences, 2017, "Single Molecule, Real-Time (SMRT)", en <http://www.pacb.com/>, consultado el 01/08/17.
- Pall, S. y A. Hamilton, 2008, "Improved northern blot method for enhanced detection of small RNA", en *Nat Protoc* 3(6): 1077-1084.
- Paterson, H. y A. Kolata, 2017, "Genomics: Keen insights from quinoa", en *Nature* 542(7641): 300-302.
- Plaza, C. *et al.*, 2017, "Impact of the 'Omics Sciences' in Medicine: New Era for Integrative Medicine", en *J Clin Microbiol Biochem Technol* 3(1): 009-0013.
- Ramírez, J., 2016, "Gen: Unidad estructural y funcional de la herencia", en L. Pereyra (comp.), *Genómica estructural y funcional en las enfermedades multifactoriales*, México, 15-51
- Rio, D., 2015, "Twenty years of RNA", en *RNA* 21(4): 718-720.
- Ruiz, G. *et al.*, 2014, "La genómica en la medicina", en *Rev Med Inst Mex Seguro Soc* 52(5): 566-573.
- Rungger, D. y M. Crippa, 1977, "The primary ribosomal DNA transcript in eukaryotes", en *Prog Biophys Mol Biol* 31(3): 247-269.
- Salleh, H. *et al.*, 2013, "A Review On Pathway Analysis Software Based On Microarray Data Interpretation", en *IJBSBT* 5(4): 149-158.
- Sandelin, A. *et al.*, 2007, "Mammalian RNA polymerase II core promoters: insights from genome-wide studies", en *Nat Rev Genet* 8(6): 424-436.
- SCOPPI. 2017, "Database of all domain-domain interactions and their interfaces derived from PDB structure files and SCOP domain definitions", en <http://scoppi.biotec.tu-dresden.de/scoppi/>, consultado 01/08/17.
- Seehausen, O. *et al.*, 2014, "Genomics and the origin of species", en *Nat Rev Genet* 15(3): 176-192.
- Sims, D. *et al.*, 2014, "Sequencing depth and coverage: key considerations in genomic analyses", en *Nat Rev Genet* 15(2): 121-132.

- Sondka, Z. *et al.*, 2017, "COSMIC Cancer Gene Census: expert descriptions across genes in oncogenesis", en *Cancer Res* 77(13 Supplement): 2599-2599.
- Stone, E. *et al.*, 2017, "Cell-selective proteomics for biological discovery", en *Curr Opin Chem Biol* 36(1): 50-57.
- STRING, 2017, "Database with information on about 9.6 million proteins from more than 2000 organisms", en <https://string-db.org/>, consultado el 01/08/17.
- Stumpf, P. *et al.*, 2008, "Estimating the size of the human interactome", en *Proc Natl Acad Sci USA* 105(19): 6959-6964.
- Sun, S. *et al.*, 2008, "Effect of fluorescently labeling protein probes on kinetics of protein-ligand reactions", en *Langmuir* 24(23): 13399-13405.
- Thermo Fisher Scientific, 2017, "GeneChip™ Human Transcriptome Array 2.0", en <https://www.thermofisher.com/>, consultado 01/08/17.
- van Dijk, L. *et al.*, 2014, "Ten years of next-generation sequencing technology", en *Trends Genet* 30(9): 418-426.
- Vogelstein, B. *et al.*, 2013, "Cancer genome landscapes", en *Science* 339(6127): 1546-1558.
- Wallden, B. *et al.*, 2015, "Development and verification of the PAM50-based Prosigna breast cancer gene signature assay", en *BMC Med Genomics* 8(1): 54-61.
- Wang, Y. y X. Qian, 2014, "Functional module identification in protein interaction networks by interaction patterns", en *Bioinformatics* 30(1): 81-93.
- Wang, Z. *et al.*, 2009, "RNA-Seq: a revolutionary tool for transcriptomics", en *Nat Rev Genet* 10(1): 57-63.
- Wray, A. *et al.*, 2003, "The evolution of transcriptional regulation in eukaryotes", en *Mol Biol Evol* 20(9): 1377-1419.
- Wrzodek, C. *et al.*, 2013, "Precise generation of systems biology models from KEGG pathways", en *BMC Syst Biol* 7(1): 15-21.

- Xu, T. *et al.*, 2016, "Identifying Cancer Subtypes from miRNA-TF-mRNA Regulatory Networks and Expression Data", en *PLoS One* 11(e): e0152792-e0152799.
- Yamamoto, T. *et al.*, 2009, "Effective Interatomic Potentials Based on The First-Principles Material Database", en *J Data Sci* 8(1): 62-69.
- Yong, M. *et al.*, 2010, "Associations between polymorphisms in glucuronidation and sulfation enzymes and mammographic breast density in premenopausal women in the United States", en *Cancer Epidemiol Biomarkers Prev* 19(2): 537-546.
- Zhang, A. y D. Trinkle, 2015, "Database optimization for empirical interatomic potential models", en *Modelling Simul Mater Sci Eng* 23(6): 1-17.
- Zhegunov, G., 2012, "Cells and Organisms", en G. Zhegunov (comp.), *The Dual Nature of Life*, 73-81. The Frontiers Collection, Springer-Verlag, Berlin, Heidelberg, Nueva York, EUA.
- Zoghbi, Y. y A. Beaudet, 2016, "Epigenetics and Human Disease", en *Cold Spring Harb Perspect Biol* 8(2): 019497-019504.

