

## Módulo Sistemas biológicos

Cuando pensamos en sistemas biológicos, se nos puede venir a la mente ecosistemas grandes y complejos con miles de especies que compiten y dependen entre ellas. Sin embargo, no necesariamente tenemos que ir tan lejos para encontrar sistemas biológicos; por ejemplo, si detallamos la maquinaria biológica de nuestras células por ejemplo los organelos que las componen, podemos observar sistemas altamente complejos, de los cuales, sólo conocemos una fracción de sus componentes. Lo anterior es debido principalmente a que los experimentos convencionales, están diseñados para deducir de forma indirecta la existencia y función de éstos componentes, así podríamos preguntarnos: ¿qué hace tan difícil estudiar los sistemas biológicos? y la respuesta no es precisamente su tamaño; por ejemplo, si observamos la red de sistema de transporte de un País podemos ver que por más grande que sea no será difícil entender su funcionamiento, conexiones para llegar a algún punto y rutas alternativas en caso de una falla en alguna estación; sin embargo si observamos una telaraña, podemos encontrar numerosos patrones dinámicos asociados a un complejo proceso de producción que aún para nuestro entendimiento es considerada como un misterio.

Los sistemas biológicos consisten de una gran cantidad de componentes, y uno de los principales retos para su análisis consiste en que los procesos que los gobiernan no son lineales. Nuestros antepasados iniciaron el estudio de los sistemas recopilando información sobre todo lo que pudiera medirse como formas, colores tamaños o patrones; estas descripciones externas aunque valiosas, no proporcionaron pistas específicas sobre su funcionamiento. La suposición implícita de este enfoque reduccionista consistía en que el estudio de los componentes básicos de la vida nos llevaría a una comprensión integral de su funcionamiento.

Avanzando hacia el siglo XXI, nos encontramos descubriendo y caracterizado genes, proteínas y metabolitos como los componentes básicos de sistemas moleculares, así, como consecuencia directa de estos esfuerzos, ese conjunto de moléculas sigue creciendo a un ritmo acelerado: con la caracterización de más de 20.000 genes

humanos, y cientos de miles de proteínas y metabolitos. Además del estudio de las partes de forma aislada, hemos avanzado en el estudio de sus relaciones y dependencia a nivel de sistema; por ejemplo: la expresión de un gen puede depender de varios factores de transcripción, metabolitos y una variedad de ARN pequeños, así como los procesos epigenéticos de su ADN.

En el siglo XIX, Claude Bernard destacó que “...la aplicación de las matemáticas a los fenómenos naturales es el objetivo de toda ciencia, porque la expresión de las leyes de los fenómenos debe ser siempre matemática...”; y un siglo después, Ludwig von Bertalanffy, trabajó intensivamente en la naturaleza sistémica de los organismos vivos. A pesar de estos esfuerzos, la biología de sistemas no tuvo un gran auge en ésta época y se mantuvo alejada de las matemáticas, la informática y la ingeniería, lo anterior debido principalmente a que los fenómenos biológicos se consideraban demasiado complicados para un análisis matemático riguroso y las matemáticas se consideraban aplicables sólo a sistemas muy pequeños de escasa relevancia biológica.

Como se mencionó anteriormente las aproximaciones reduccionistas son muy útiles en la generación de información detallada sobre componentes o procesos específicos de un sistema; sin embargo carecen de la capacidad de caracterizar, explicar o predecir propiedades emergentes del estudio integrado de los componentes de un sistema; por ejemplo hace unos años, el laboratorio de Hirotada Mori completó el ensamblaje de un catálogo completo de mutantes únicos en *E coli*. Sin embargo, aún no podemos saber qué genes están prendidos o apagados en respuesta a condiciones ambientales variantes.

Otro ejemplo muy desafiante de las propiedades del sistema emergente es el sistema nervioso central. Aunque comprendemos bastante bien cómo se generan y propagan los potenciales de acción en las neuronas individuales, aún no entendemos en su totalidad cómo fluye la información, cómo funciona la memoria y cómo las enfermedades afectan el funcionamiento normal del cerebro. Ni siquiera está claro cómo se representa la información en el cerebro. Por lo tanto, si bien la biología

reduccionista ha tenido un gran éxito y, sin duda alguna, seguirá siendo la principal fuerza impulsora para futuros descubrimientos, muchos científicos hemos llegado a reconocer que la información detallada resultante de este enfoque debe complementarse con nuevos métodos a nivel de sistemas e integración de información.

Lo anterior podemos entenderlo mejor retomando un cuento indio que describe a seis ciegos explorando un elefante. Esta historia es bastante antigua y generalmente termina en una confusión total, pero es útil analizarla un poco más allá: cuenta la historia que cada uno de los ciegos tocó una parte diferente de un elefante y llegó a una conclusión diferente con respecto al objeto de su investigación. El hombre que tocaba el costado pensó que estaba tocando una pared, el que sentía la pierna concluyó que estaba tocando un árbol. La trompa del elefante daba la impresión de una serpiente, el colmillo el de una espada puntiaguda, la cola se sentía como una cuerda y la oreja parecía una gran hoja o abanico.

En las últimas décadas, se ha establecido una sólida base teórica y numérica para los métodos computacionales diseñados específicamente para la investigación de sistemas. Los avances algorítmicos permiten la simulación y optimización de redes de distribución de flujo biológico muy grandes; algunos de éstos métodos están comenzando a hacer posible la predicción con un alto nivel de confianza del plegamiento de proteínas; éstas predicciones, a su vez, permiten la caracterización de interacciones moleculares específicas, por lo que son actualmente un blanco de atención como un potencial de intervenciones farmacológicas dirigidas hacia la minimización de efectos secundarios tóxicos.

No es difícil ver la analogía con un sistema biológico complejo como el inicio de la enfermedad de Alzheimer. El primer científico encontró "el gen de Alzheimer", el segundo descubrió "una fuerte asociación entre la enfermedad y lesiones o traumas", otro científico detectó "problemas con el metabolismo de los ácidos grasos en el cerebro" y otro sugirió que "el aluminio de ollas de cocina podría ser el culpable". Como en el caso del elefante, los científicos tenían razón, hasta cierto punto.

La biología de sistemas no les pide a los seis ciegos que abandonen sus métodos, al centrarse en un aspecto, ya que son expertos en la parte de su cuerpo elegida en cambio, lo que sugiere es, la integración de otros investigadores y disciplinas desarrollando nuevas herramientas de análisis, diciéndoles en su idioma lo que otros han encontrado y así cerrando la brecha entre troncos y colmillos y colas.

Una estrategia para lograr esta sinergia es la recopilación de los diversos datos e información contextual obtenidos por los seis ciegos y fusionarlos en un modelo conceptual. ¿Qué tipo de “cosa” podría consistir en partes que parezcan una serpiente, troncos de árboles, paredes grandes, dos espadas, dos abanicos y una cuerda? ¿Cómo podrían interactuar funcionalmente un grupo de genes, trauma craneoencefálico y un metabolismo cerebral alterado para producir la enfermedad de Alzheimer?

Los métodos para estos análisis se pueden automatizar y aplicar utilizando herramientas computacionales. Los resultados equivalen por ejemplo a patrones de conectividad en las redes biológicas estudiadas sugiriendo así roles importantes de estos componentes en diferentes escenarios. El análisis también puede mostrar qué variables actúan entre sí y si el patrón de conectividad dentro de la red cambia, por ejemplo, durante el desarrollo y el envejecimiento o en respuesta a perturbaciones, como una enfermedad.

Muchos conceptos y métodos de análisis de redes son en realidad bastante simples e intuitivos cuando se estudian en redes pequeñas. Esta simplicidad es crucial para el análisis de redes, es decir, su escalabilidad a conjuntos de moléculas muy grandes y su capacidad para revelar patrones que la mente humana sin ayuda simplemente no sería capaz de reconocer. Por lo tanto, los programas pueden entrenarse para usar los métodos básicos una y otra vez, establecer modelos de red a partir de datos experimentales, identificar sus características globales y locales y, en última instancia, conducir a hipótesis sobre su funcionalidad.

Los modelos mayormente estudiados son estáticos; siendo una aproximación válida en la medida en que los cambios en componentes de interés suelen ser relativamente lentos. Por ejemplo, incluso si un organismo está creciendo, su tamaño es casi constante durante un experimento bioquímico que dura unos minutos. En segundo lugar, se puede aprender mucho asumiendo que una red es al menos temporalmente estática. Un ejemplo típico puede ser el perfil de actividad de las enzimas en el hígado, que cambia a lo largo de la vida de un organismo e incluso dentro de un período de 24 horas, pero que puede ser más o menos constante en un período de unos pocos segundos o minutos. Sumado a lo anterior los estados estacionarios suelen ser más fáciles de caracterizar que los cambios dinámicos y, una vez establecidos, pueden proporcionar un punto de partida importante para examinar las características dinámicas de un sistema

El análisis de éstas redes biológicas se puede subdividir de varias formas, esto permite también diferentes métodos de análisis. En algunos casos, por ejemplo, podemos considerar el flujo de una red y cuantificar cantidades o masas de reacciones así en el estudio de rutas metabólicas podríamos suponer la cantidad de producto a partir de una reacción donde conocemos el sustrato consumido; éste análisis de equivalencia es muy útil para los análisis matemáticos. De acuerdo con el tipo de datos se pueden generar diferentes redes en términos de conectividad, complejidad y estructura.

Algunos de los tipos más comunes de redes son: Redes de interacción proteína-proteína, Redes metabólicas, Redes de interacción genética, Redes de regulación génica / transcripcionales, Redes de señalización celular entre otras.

Un segundo criterio en el estudio de redes equivale al tipo de investigación realizada; donde podríamos considerar el análisis y/o la reconstrucción de redes. Un análisis típico asume que la estructura de una red ya es conocida e investiga las características como la conectividad entre sus componentes; por otro lado una reconstrucción implica inferir la conectividad de una red desconocida o no caracterizada a partir de datos experimentales y literatura. En cualquiera de los dos

casos la interpretación de los componentes de la red es secundaria; es decir, para el análisis matemático y computacional no es relevante si sus componentes son moléculas, especies o subpoblaciones porque todos ellos son considerados como nodos conectados por aristas. Las aristas pueden ser dirigidas, no dirigidas o bidireccionales.

En conjunto los nodos y las aristas forman grafos, los cuales, son las representaciones de las redes, de una forma intuitiva podemos observar que los nodos de una red pueden estar sólo conectados por aristas y dos aristas pueden sólo estar conectadas por un nodo entre ellas. 2 nodos conectados por una sola arista son llamados nodos vecinos. Algunos ejemplos típicos de grafos dirigidos son las redes de regulación transcripcional donde los nodos representan los genes y las aristas denotan interacciones entre ellos.

La subdisciplina de las matemáticas y ciencias computacionales que analizan las redes estáticas se conoce como teoría de grafos. De forma general la teoría de grafos, se ocupa de las propiedades de los grafos como su conectividad y rutas más cortas entre dos nodos o rutas que conectan todos los nodos. Existen muchos problemas matemáticos famosos que están asociados con grafos, como el teorema de 4 colores, que postula que siempre es posible colorear un mapa geográfico con 4 colores de tal forma que las ciudades en un mismo país nunca tendrían el mismo color. Otro problema famoso que combina grafos y optimización es el agente de ventas que busca la ruta más corta en un país para visitar una ciudad a la vez en el recorrido.

Este enfoque para el estudio de los sistemas considera entonces la biología como una ciencia de la información. Las nuevas tecnologías de alto rendimiento y algoritmos de visualización y modelamiento han transformado la biología haciendo posible la medición y cuantificación de información biológica de diferentes tipos como secuencias de genomas completos, identificación y cuantificación de comunidades de microbioma, medición de niveles de expresión de genes, proteínas, metabolitos además ha permitido la integración de información a diferentes niveles, abonando

esfuerzos para avanzar en el descubrimiento de biomarcadores y métodos de diagnóstico temprano de enfermedades.

Los métodos diagnósticos tradicionales, usualmente se han centrado en un parámetro o un conjunto reducido de parámetros para descifrar una enfermedad (e.g. un marcador conocido en el diagnóstico de cáncer de próstata es un antígeno específico); así la habilidad para diferenciar de forma exacta entre salud y enfermedad o identificar diferentes categorías o subtipos de la enfermedad ha sido extremadamente limitada. La aproximación de estudio de los fenómenos biológicos a través de la biología de sistemas permitirá el crecimiento de una nueva disciplina médica conocida como la medicina de sistemas- que ha comenzado a surgir recientemente. La premisa central de la medicina de sistemas es que las firmas moleculares clínicamente detectables resultan a partir de la perturbación en las redes biológicas causada por condiciones patológicas. Los procesos de enfermedad asociados a las firmas moleculares podrá ser usado para agrupar a individuos en sub-poblaciones basado en variantes genéticas que afectan la progresión de la enfermedad.

La clave de esta aproximación radica en el uso de los modelos de redes biológicas, combinados con la información molecular obtenida a partir de muestras de pacientes. Los rápidos avances en las tecnologías de alto rendimiento, incluida la secuenciación del genoma, ciencias ómicas, nuevos métodos de visualización y el uso de tecnologías computacionales, están cambiando los métodos diagnósticos actuales. Las "firmas moleculares" asociadas a procesos patológicos específicos (panel de biomarcadores) pueden estar compuestas por diferentes tipos de biomoléculas, incluidas proteínas [1], ADN [2], ARN [3] o microARN (miARN) [4] y metabolitos [5], así como modificaciones postraduccionales de estas moléculas, como la fosforilación de proteínas [6] y la glicosilación [7,8]. Los análisis precisos de múltiples parámetros serán la clave para identificar, evaluar y rastrear estos patrones moleculares que reflejan redes perturbadas por enfermedades.

La pregunta esencial de cómo la biología de sistemas facilita el descubrimiento de biomarcadores informativos se podría responder a partir de los siguientes puntos

clave: 1. Las redes biológicas presentan perturbaciones de su estado normal en un proceso de enfermedad, 2. Los componentes Órgano-específicos de estas redes pueden ser usados para inferir el estado de redes específicas de tejidos blanco 3. Los paneles de múltiples marcadores sanguíneos pueden aportar abundante información sobre lo que está ocurriendo tipos celulares específicos 4. Las proteínas o RNAs liberados en la sangre proporcionan una evaluación del estado y evolución futura de redes perturbadas. El futuro de la aproximación basada en medicina de sistemas se encuentra basada en la habilidad para identificar y evaluar el estado de las redes perturbadas, permitiendo así tomar decisiones en cuanto a intervenciones adecuadas.

El conocimiento de la dinámica de la red y sus interacción nos permitirá avanzar en la predicción de la progresión de la patología basado en el estudio de sus interacciones, así mismo abrirá una puerta a intervenciones específicas o modificaciones de subredes clave para prevenir la evolución predicha de la patología.

El futuro de la medicina de sistemas consistirá en la aplicación de nuestro entendimiento a las respuesta de redes moleculares para determinar estados críticos de la enfermedad para evaluar, corregir, modificar o ajustar las respuestas de la red para revertir el estado patológico o prevenir la progresión de la enfermedad a un estado más severo. Este futuro aún está fuera de nuestro alcance, pero la ruta para alcanzarlo es cada vez más clara.

## **Referencias**

1. X. Wang et al., Autoantibody signatures in prostate cancer, N Engl J Med 353(12) (2005), 1224–1235.
2. E. Papadopoulou et al., Cell-free DNA and RNA in plasma as a new molecular marker for prostate and breast cancer, Ann N.Y. Acad Sci 1075 (2006), 235–243.



3. K.D. Sorensen and T.F. Orntoft, Discovery of prostate cancer biomarkers by microarray gene expression profiling, *Expert Rev Mol Diagn* 10(1), 49–64.
4. C.R. Scherzer et al., Molecular markers of early Parkinson's disease based on gene expression in blood, *Proc Natl Acad Sci USA*, 2007.
5. K.S. Solanky et al., Application of biofluid <sup>1</sup>H nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile, *Anal Biochem* 323(2) (2003), 197–204.
6. A. Deguchi et al., Vasodilator-stimulated phosphoprotein (VASP) phosphorylation provides a biomarker for the action of exisulind and related agents that activate protein kinase G, *Mol Cancer Ther* 1(10) (2002), 803–809.
7. P.A. Norton et al., N-linked glycosylation of the liver cancer biomarker GP73, *J Cell Biochem* 104(1) (2008), 136–149
8. P.M. Drake et al., Sweetening the Pot: Adding Glycosylation to the Biomarker Discovery Equation, *Clin Chem*, 2009.







