



La “materia oscura” de los sistemas biológicos

Juan Antonio García Ranea

Investigador Ramón y Cajal. Departamento de Biología Molecular y Bioquímica. Universidad de Málaga. ranea@uma.es

La materia oscura es un concepto surgido en la Astrofísica para definir aquella masa del universo que no podemos ver con los medios técnicos actuales, pero cuya existencia predecimos por la influencia indirecta que ejerce en el devenir de estrellas y galaxias.

En la era de la biología de sistemas hemos comenzado nuestra andadura en la investigación del cosmos de interacciones moleculares que articulan los sistemas vivos. En este contexto, y de forma análoga a lo ocurrido en la exploración del universo, los científicos comenzamos a intuir que grandes espacios del intrincado mundo de interrelaciones moleculares subyacentes a la vida aún permanecen oscuros a nuestro conocimiento.

La irrupción de la “revolución industrial” en los experimentos de biología molecular nos ha hecho creer que por fin contamos con los medios técnicos necesarios para la exploración de estos vastos espacios desconocidos. Pero la gran paradoja es que a pesar de que contamos con toneladas de datos experimentales no somos capaces de enunciar modelos sistémicos consistentes que expliquen la diversidad funcional de los sistemas biológicos complejos. Ante esta paradoja, cabría preguntarse ¿Cuál es la razón o razones que, en biología de sistemas, limitan tan significativamente la transformación de la información en conocimiento?

De forma general, la comunidad científica ha apuntado como principal sospechoso de dicha frustración a la falta de un mayor desarrollo formal y conceptual en una disciplina incipiente como es la Biología de Sistemas. No obstante, y sin menoscabo de que esta primera razón sea cierta, cabría también cuestionarse si con el conocimiento acumulado hasta la fecha, más las técnicas experimentales de exploración actuales, estamos en el buen camino para llegar a comprender como funcionan los sistemas moleculares que estudiamos. En resumen, podríamos decir que mientras la primera razón tendría que ver con una insuficiencia del desarrollo teórico de la biología de sistemas, la segunda apuntaría a un desarrollo ineficiente de la exploración experimental del espacio de interacciones, generándose así modelos incompletos difíciles de interpretar.

Adivinar qué progreso teórico futuro terminará por salvar las limitaciones actuales en este área de conocimiento se presenta como una tarea más especulativa que científica. En cambio, con los medios computacionales actuales sí es factible estimar el efecto que nuestras técnicas experimentales ejercen en el modelado de ciertos sistemas moleculares importantes, como por ejemplo lo son las redes de interacción entre proteínas. Esto es posible porque, al igual que en Astronomía la materia oscura se predice por su influencia indirecta en otros cuerpos celestes, muchas interacciones de las que aún no disponemos de evidencia experimental son producto de largos procesos de co-dependencia evolutiva entre proteínas que dejan su impronta en la estructura y función de los genomas, señales que son detectables mediante distintos métodos bio-computacionales.

Con el objeto de estudiar los males y virtudes de nuestras técnicas experimentales de exploración y estimar la suficiencia de los datos acumulados por la comunidad científica durante décadas, los grupos de la Profesora Christine Orengo, en la *University College London*, y nuestro grupo, en la Universidad de Málaga, colaboran en un estudio comparativo entre la red de interacciones de proteínas basada en el compendio de todas las evidencias experimentales disponibles en bases de datos públicas, y la red de interacciones predichas mediante métodos bio-computacionales [1]. Para dicho estudio utilizamos como objetos de investigación los proteomas completos de dos especies eucariotas muy diferentes, como son el hombre y la levadura *S. cerevisiae*. Especies que se consideran entre las mejor caracterizadas desde el punto de vista de la Biología Molecular entre los eucariotas, y a la vez lo suficientemente alejadas en el árbol de la vida como para poder ofrecernos la posibilidad de encontrar patrones universales en nuestros estudios. De esta forma

tenemos dos modelos independientes de la red de interacciones de proteínas que comparan en humano y levadura, por un lado una red compendio del conocimiento derivado de datos experimentales (red experimental), y por otro, la red de interacciones predicha mediante técnicas bio-computacionales (red de predicciones).

Con la comparación y estudio de ambas redes tratamos de dar respuestas a las siguientes preguntas: ¿Cuánto espacio de las redes de interacciones de proteínas permanece aún oscuro a nuestro conocimiento experimental? ¿Este espacio oscuro aún por explorar encierra componentes funcionalmente claves necesarios para comprender el funcionamiento del sistema? ¿Nuestras técnicas actuales de exploración experimental nos permitirán con el tiempo caracterizar todos los rincones del espacio de interacciones, o presentan limitaciones de muestreo que están dejando en la oscuridad importantes regiones funcionales?

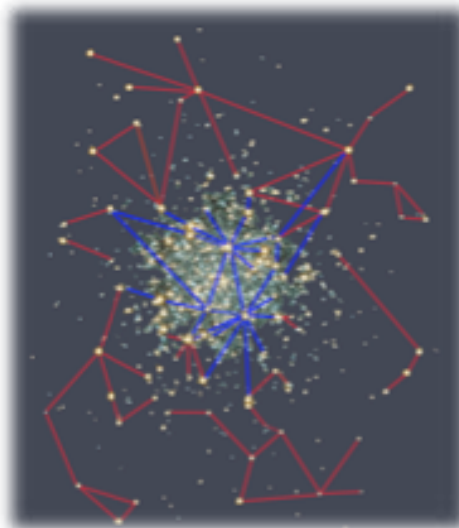
El estudio de solapamiento entre ambas redes [1] nos desveló que una proporción muy significativa de las redes de interacción entre proteínas predichas por medios bio-computacionales no ha sido aún caracterizada experimentalmente (el 82% en levadura y el 98% en humanos). A este espacio de interacción inexplorado lo denominamos la “materia oscura” de las redes de interacción. Aunque un alto porcentaje de materia oscura fue observado en ambos organismos, en la levadura el porcentaje de predicciones de las que sí tenemos evidencia experimental (18%) se estimó del orden de 14 veces superior que en humanos (1,34%), indicándonos que al menos desde el punto de vista de la suficiencia de los datos experimentales actuales, probablemente nos encontramos mucho más cerca de comprender el funcionamiento de los sistemas en *S. cerevisiae* de lo que estamos en el ser humano.

En la búsqueda de componentes funcionalmente claves en estos espacios no caracterizados experimentalmente, encontramos un alto número de proteínas *hubs* o nodos centrales de conexión, claves en la arquitectura de relaciones de la red. Al estudiar el papel funcional de estas proteínas que denominamos “*hubs* ocultos”, descubrimos que en la levadura correspondían principalmente a proteínas de membrana, mientras que en humanos estaban relacionados con funciones regulatorias en vías de señalización, como

las señales mediadas por proteínas quinasas mediante interacciones transitorias y que están implicadas en multitud de procesos de transmisión de señales y regulación celulares fundamentales de los organismos eucariotas.

Los resultados obtenidos en estos experimentos ponen en cuestión la manera en la que exploramos experimentalmente las redes de interacción molecular, así como nuestras posibilidades de llegar a modelar y comprender su funcionamiento. La estimación de la materia oscura nos indica que incluso para los organismos mejor estudiados como son el hombre y la levadura, en estos momentos estamos muy lejos de disponer de un catálogo suficiente de componentes y relaciones que nos permita generar modelos completos. Además, no solo nos falta mucha información, sino que desconocemos muchos componentes clave en el funcionamiento de dichas redes, como lo demuestra la existencia de nodos centrales o *hubs* ocultos. Por tanto, con datos incompletos e información determinante no disponible, es obvio concluir que nos encontramos muy lejos de poder generar modelos consistentes que nos permitan reproducir y entender el funcionamiento de dichos sistemas complejos [1,2].

La insuficiencia de datos no sería en sí demasiado preocupante si supiéramos que es sólo cuestión de tiempo que la aplicación recurrente de las técnicas experimentales actuales termine la caracterización de dichos espacios oscuros. Pero desafortunadamente, el análisis funcional de la materia oscura nos indica que nuestra exploración experimental de este espacio es sesgada, ignorando de forma sistemática regiones funcionalmente fundamentales, sin las cuales difícilmente alcanzaremos



a comprender el funcionamiento de los sistemas moleculares. Todos conocemos las limitaciones y el bajo rendimiento que distintas técnicas experimentales presentan en la caracterización, por ejemplo, de proteínas de membrana, las cuales están subrepresentadas en las bases de datos actuales, o las limitaciones que métodos como el sistema de doble híbrido en levadura presentan en la detección de interacciones transitorias como, por ejemplo, las que se producen en vías de señalización y cascadas de regulación protagonizadas por las proteínas quinasas tan abundantes en los proteomas de eucariotas complejos [3].

A la luz de los resultados actuales, podemos concluir que el avance del conocimiento en biología de sistemas no solo depende del desarrollo de nuevos conceptos, sino también de nuevos diseños experimentales de alto rendimiento que nos permitan corregir y compensar el sesgo de las técnicas aplicadas actualmente, así como completar la caracterización de las relaciones funcionales que permanecen oscuras y de las que por ahora sólo podemos tener evidencias indirectas.

La cosmovisión del hombre ha sufrido cambios dramáticos a medida que su conocimiento del universo se ha hecho más preciso, pasando de

un modelo geocéntrico a otro heliocéntrico, y posteriormente a la comprobación de la humilde ubicación de nuestro Sistema Solar en un brazo lateral de la galaxia Vía Láctea. Es probable que, de igual forma, se tenga que vencer la resistencia cultural a revisar alguno de los paradigmas y dogmas más queridos por parte de comunidades y grupos científicos bien establecidos en el ámbito de la Biología, lo que implicaría aceptar la idea de que probablemente las proteínas que se piensan centrales en muchos de los modelos experimentales actuales deban compartir su protagonismo con otras proteínas cuyo papel central no ha sido descubierto todavía. E incluso, en algunos casos, proteínas ahora consideradas claves deban dejar su papel protagonista a otras en modelos posteriores más precisos a medida que nos adentremos en la "materia oscura" de los organismos vivos.

Bibliografía citada:

1. Ranea JA, Morilla I, Lees JG, Reid AJ, Yeats C, Clegg AB, Sanchez-Jimenez F, Orengo C (2010). Finding the "dark matter" in human and yeast protein network prediction and modelling. *PLoS Comput Biol.* 6: pii: e1000945.
2. Wass MN, David A, Sternberg MJ (2011). Challenges for the prediction of macromolecular interactions. *Curr Opin Struct Biol.* 21: 382-90.
3. Russell RB, Aloy P (2008). Targeting and tinkering with interaction networks. *Nat Chem Biol* 4: 666–673.

