

12 DE FEBRERO DE 2001

## El análisis del genoma humano revela nuevas proteínas involucradas en la expresión génica

Una búsqueda inicial en el bosquejo de la secuencia del genoma humano ha revelado evidencias prometedoras que sugieren que el genoma completo revelará nuevas proteínas involucradas en la expresión génica. Los investigadores que realizaron el análisis descubrieron que genes previamente desconocidos parecen codificar para proteínas involucradas en la expresión génica.

Los investigadores también encontraron evidencias de que la maquinaria humana de expresión génica es más compleja que la de animales inferiores. Este descubrimiento sugiere que una maquinaria más sofisticada para traducir la información genética "puede ser particularmente importante para el desarrollo y la fisiología humana", escribieron los investigadores.

---

"Esta búsqueda ilustra cómo la secuencia del genoma humano proporcionará muchos factores nuevos que pueden estar involucrados en la expresión génica, aunque sus funciones siguen siendo desconocidas."

— **Michael R. Green**

---

El investigador del Instituto Médico Howard Hughes, Michael R. Green, que se encuentra en la Facultad de Medicina de la Universidad de Massachusetts y los coautores Rossella Tupler de la Università degli Studi di Pavia y Giovanni Perini de la Universidad de Bolonia publicaron su análisis en el número del 15 de febrero de 2001, de la revista *Nature*. El análisis es parte de una colección de artículos publicados por *Nature*, que discuten las implicancias de la labor de secuenciación del genoma humano.

"La disponibilidad del genoma humano y de las secuencias de otros genomas revolucionará todos los campos de investigación biomédica", escribieron los científicos. "Pero, dado que el genoma mismo es objeto de la expresión génica, el impacto puede ser particularmente importante para los que estudiamos este proceso".

En una entrevista en la que discute el artículo de *Nature*, Green agrega, "en el estudio de la expresión génica, el genoma mismo es el que está siendo explorado, a diferencia de lo que ocurre en la mayoría de los otros procesos biológicos. El genoma contiene las secuencias de los jugadores involucrados—las proteínas—y las secuencias de todas las señales que gobiernan la expresión génica".

En los análisis, los científicos buscaron los genes que gobiernan tres pasos de la expresión génica—la transcripción del gen a ARN mensajero (ARNm); la maduración por corte y empalme inicial del ARNm, para producir la molécula final que especificará una proteína; y la adición de la "cola" de poli(A) en un extremo del ARNm, que es necesaria para su procesamiento por la maquinaria de producción de proteínas.

Cuando buscaban en el genoma humano las secuencias génicas que especifican factores generales de transcripción (GTFs, por sus siglas en inglés) y activadores transcriptionales, los investigadores encontraron las secuencias de numerosos genes que eran similares a los de levadura y de la mosca de la fruta *Drosophila*. Sin embargo, también encontraron muchas más secuencias de genes humanos que en el genoma de *Drosophila*, que parecían estar relacionadas con un GTP particular, llamado TFIID, "lo que indica que la diversidad potencial del TFIID humano es mucho mayor que la diversidad del de *Drosophila*".

Al buscar activadores transcripcionales, los investigadores encontraron más de 2.000 genes que podían codificar para estas proteínas—muchos más que los encontrados en las bases de datos del genoma de *Drosophila* y del gusano redondo *C. elegans*. "Esta búsqueda ilustra cómo la secuencia del genoma humano proporcionará muchos factores nuevos que pueden estar involucrados en la expresión génica, aunque sus funciones siguen siendo desconocidas", escribieron los científicos.

De la misma forma, en una búsqueda para los genes que codifican para las proteínas implicadas en ensamblar la maquinaria de maduración por corte y empalme del ARNm, los científicos encontraron una "complejidad significativamente mayor que la que se encuentra en *Drosophila*". Y al buscar los genes que codifican para las proteínas implicadas en la adición de la cola de poli(A) al ARNm, los investigadores encontraron inesperadamente nuevos genes, lo que nuevamente sugiere que los seres humanos tienen una maquinaria más compleja de expresión génica.

Green dice, "en cada caso, las sorpresas son del mismo tipo. Cuando buscábamos los genes que codificaban para ciertos tipos de proteínas que eran homólogas a las de *Drosophila*, no teníamos motivos para creer que habría más de una sola proteína. Pero, de hecho, encontramos homólogos para los que no teníamos motivos para sospechar que estarían allí".

Los investigadores señalan, sin embargo, que "aunque estas búsquedas destacan el poder de la nueva información genómica, también revelan limitaciones importantes. En particular, la existencia de un gen relacionado no significa que haya una proteína correspondiente: la secuencia podría

corresponder a un pseudogén que no es expresado".

Green explicó que los pseudogenes son "fósiles" de genes verdaderos. "Pero el genoma no los ha expresado", dijo. "Los ha mantenido allí, pero no hacen nada".

"Sospecho que muchos de estos nuevos genes no son pseudogenes. En algunos casos podemos encontrarlos en las bases de datos de genes expresados. Y están conservados en un grado mayor del que se predeciría si fueran pseudogenes. Si simplemente no se expresaran, no habría presión evolutiva para mantener la conservación de la secuencia".

Los resultados plantean una variedad de preguntas intrigantes que tomará tiempo responder, dice Green. "Estos resultados son muy interesantes y emocionantes, pero, por supuesto, ahora tenemos que empezar a averiguar su significado en el contexto de la expresión génica. ¿Estas proteínas adicionales se originan durante ciertas etapas del desarrollo? ¿En ciertas células? ¿Qué hacen? ¿Cuáles son sus genes diana? Pero no estaríamos en posición de hacernos estas preguntas si no conociéramos que estos candidatos existen".

Muchos de los genes nuevos parecen expresar subunidades de complejos proteicos más grandes. Estudios adicionales serán necesarios para entender cómo están involucrados en esos complejos. "Tal cual sucede con cualquier descubrimiento, este hallazgo abre toda una serie de nuevos interrogantes y metodologías que se pueden explorar", dijo Green.