

SEBBM DIVULGACIÓN

ACÉRCATE A NUESTROS CIENTÍFICOS



El genoma regulador

DOI: http://dx.doi.org/10.18567/sebbmdiv_ANC.2018.04.1

José Luis Gómez Skarmeta

Centro Andaluz de Biología del Desarrollo, CSIC-Universidad Pablo de Olavide, Sevilla

Biografía

José Luis Gómez Skarmeta, Profesor de Investigación del CSIC en el Centro Andaluz de Biología del Desarrollo (Sevilla), es licenciado en Químicas por la Universidad de Murcia (1984) y doctorado en Ciencias por la Universidad Autónoma de Madrid (1995). Sus áreas de especialidad son la Biología del Desarrollo y la Evolución, la Biología Molecular, la Genética, la Genómica Funcional y la Epigenómica. Durante su carrera investigadora ha publicado más de 100 artículos en prestigiosas revistas de investigación. El Dr. Gómez Skarmeta ha sido pionero en combinar novedosas técnicas de epigenómica y de biología del desarrollo para estudiar la contribución de la información reguladora y la estructura de la cromatina al desarrollo, la evolución y las enfermedades humanas. Entre sus principales aportaciones recientes están la identificación del gen *Irx3* como gen diana de las principales mutaciones asociadas a obesidad (Smemo y cols., *Nature* 2014), los estudios que demuestran la reorganización tridimensional del complejo *Hox* en la transición invertebrado-vertebrado (Acemel y cols. *Nature Genetics* 2016) y el reciente descubrimiento sobre la aleta dorsal de peces como origen evolutivo de las extremidades de vertebrados (Letelier y cols. *Nature Genetics* 2018).

Resumen

El 98% del genoma es no-codificante y hasta hace poco se consideraba ADN basura. Los recientes avances en genómica han demostrado su función esencial en la regulación de la expresión génica, y su papel clave en la construcción de los organismos, la evolución de las especies y en las enfermedades humanas.

Summary

98% of the genome is non-coding and until recently it was considered junk DNA. Recent advances in genomics have demonstrated its essential function in the regulation of gene expression, and its key role in the construction of organisms, the evolution of species and in human diseases.

<http://www.sebbm.es/>

HEMEROTECA: http://www.sebbm.es/ES/divulgacion-ciencia-para-todos_10/acercate-a-nuestros-cientificos_107

En los genomas de vertebrados sólo una pequeña fracción de ADN se corresponde con secuencias codificantes. Además, la gran mayoría de las proteínas codificadas por estos genes están altamente conservadas en todas las especies. ¿Cómo ha surgido entonces la diversidad morfológica durante la evolución? Actualmente, la teoría genética de la evolución morfológica propone que las variaciones morfológicas entre especies se deben, en gran medida, a la alteración en la expresión de genes durante el desarrollo embrionario (Carroll 2008). La expresión de los genes está controlada por secuencias reguladoras localizadas en el ADN no-codificante que rodea dichos genes. Así, son las variaciones en estas regiones reguladoras las principales fuentes de variación morfológica.

La regulación de los genes no sólo es crucial para el desarrollo, sino también es esencial para el desarrollo embrionario y la fisiología de las células en los organismos adultos. Por lo tanto, no es sorprendente que la gran mayoría de las lesiones asociadas a enfermedades genéticas humanas se localicen en el ADN regulador (Zhang 2015). A pesar de la evidencia cada vez mayor de la función fundamental del ADN regulador en el desarrollo, la evolución y las enfermedades humanas, hasta hace poco teníamos un gran problema: mientras que el código de las secuencias codificantes de las proteínas se conoce, el lenguaje del ADN regulador es en gran medida desconocido, lo que impedía la identificación de las regiones reguladoras en los genomas secuenciados. El desarrollo de técnicas basadas en la secuenciación masiva ha cambiado completamente este panorama y hoy no sólo podemos identificarlas, sino que además hemos descubierto la importancia de su organización tridimensional en el núcleo para su correcto funcionamiento.

Varias han sido las iniciativas para identificar todos los elementos funcionales en la secuencia del genoma humano. Así, el proyecto ENCODE (Enciclopedia de Elementos de ADN), liderado por el National Human Genome Research Institute (NHGRI), pretende identificar toda la actividad funcional del genoma humano, utilizando como material multitud de líneas celulares humanas. Otro proyecto con objetivos similares, pero usando diferentes tejidos humanos tanto de embriones como de adultos es el Roadmap Epigenomics, también liderado por el gobierno de EE.UU. a través del Instituto Nacional de Salud (NIH). Estos proyectos se han complementado con otros financiados por el mismo organismo para estudiar también los genomas de modelos animales como el gusano (*Caenorhabditis elegans*), la mosca del vinagre (*Drosophila melanogaster*) y ratones (Mouse Encode y Mouse Encode). En todos estos proyectos se han combinado diferentes experimentos de secuenciación masiva para identificar a nivel de todo el genoma la cromatina abierta, las regiones donde se localizan diferentes modificaciones de histonas, los sitios de unión en el genoma de diferentes factores de transcripción y los estados de metilación del DNA.

Todos estos trabajos, junto con otros realizados en otras especies de vertebrados e invertebrados en numerosos grupos de investigación de multitud de países, han permitido identificar marcas epigenómicas universales asociadas a estados específicos de la cromatina. Por ejemplo, la trimetilación de la lisina 4 de la Histona 3 (H3K4me3), es una marca epigenómica característica de promotores activos, mientras que la acetilación de la lisina 27 de la misma histona (H3K27ac), se asocia a elementos reguladores activos. Además, se ha visto que tanto promotores como elementos reguladores activos se localizan en regiones abiertas del genoma (Shlyueva 2014; Figura 1).

A través de estos estudios se ha revelado la enorme cantidad de información reguladora en los genomas de los animales, así como su gran dinámica tanto en el tiempo como en el espacio. Es de hecho ese dinamismo espacial y temporal lo que permite expresar de forma diferencial los diferentes genes del genoma y generar así diferentes tipos celulares y órganos. Con estos principios universales se abre ahora la posibilidad de estudiar la información reguladora de cualquier genoma. Esto permitirá descifrar cómo los cambios en dicha información contribuyen a los cambios morfológicos durante la evolución.

¿Cómo se organiza esta gran cantidad de información reguladora en el genoma para que las diferentes regiones actúen sobre su gen diana y no sobre otro promotor no relacionado?

El desarrollo de técnicas basadas también en secuenciación masiva que permiten detectar contactos ADN-ADN en todo el genoma, denominadas técnicas de captura de conformación cromosómica, ha revolucionado nuestra visión sobre la organización 3D del genoma. Así, se ha demostrado que la cromatina está formando compartimentos denominados dominios topológicamente asociados (TAD). Estas son regiones genómicas del orden de una megabase en las cuales las secuencias de ADN interaccionan preferentemente entre sí. Los TAD actúan como unidades organizativas básicas del genoma, compartimentando la información reguladora en el espacio tridimensional del núcleo, generando fronteras que separan funcionalmente las regiones genómicas que están muy cerca de una perspectiva lineal y que unen las secuencias de interacción funcional que están separadas por cientos de kilobases. De esta forma elementos reguladores y sus genes dianas se localizan dentro del mismo TAD y se separan del siguiente que contiene otros genes y sus correspondientes secuencias reguladoras (Rowley 2016, Figura 1).

La combinación de los análisis de la estructura 3D de la cromatina con la detección de regiones reguladoras nos permite un conocimiento sin precedentes de la complejidad del aparato regulador que controla la expresión espacio-temporal de cualquier gen y de cómo alteraciones en este aparato pueden causar enfermedades o pueden contribuir a la evolución de las

especies. Así, estudios recientes han podido demostrar cómo los reordenamientos cromosómicos que alteran la estructura 3D de la cromatina causan la expresión génica ectópica en diferentes tejidos humanos, generando malformaciones o causando patologías como cáncer (Lupiañez 2016). Esta sobreexpresión se debe a que el reordenamiento genera un nuevo TAD y determinados genes se activan por elementos reguladores que en condiciones normales estarían en otro compartimento del genoma. Es de esperar que este mismo tipo de reordenamientos también contribuyan a importantes cambios de expresión a lo largo de la evolución que favorezcan modificaciones morfológicas, como de hecho hemos podido demostrar en estudios recientes en nuestro laboratorio (Acemel 2017).

Tal es la importancia de la organización 3D del genoma en la regulación de la expresión génica, que el NIH ha arrancado un nuevo proyecto, denominado 4D Nucleome, para estudiar este tema en profundidad.

Cabe esperar que en próximos años podamos determinar con mucha mayor precisión el impacto de las mutaciones regulatorias en la expresión génica y en el desarrollo de enfermedades genéticas humanas, así como su papel en la evolución de las especies. En gran medida, gracias al apoyo determinante de gobiernos con visión de futuro. El desarrollo económico derivado de este conocimiento será inmenso. En España, estaremos en la retaguardia, y nos va a salir caro, sobre todo para las arcas públicas. Eso nos pasa por dejar que investiguen otros.

Referencias:

1. Carroll SB. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, 34:25-36, 2008.
2. Zhang F, Lupski JR. Non-coding genetic variants in human disease. *Hum Mol Genet.* 24:R102-10, 2015.
3. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet.* 15:272-8, 2014.
4. Rowley MJ, Corces VG. The three-dimensional genome: principles and roles of long-distance interactions. *Curr Opin Cell Biol.* 40:8-14, 2016.
5. Lupiañez DG, Spielmann M, Mundlos S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet.* 32:225-237, 2016.
6. Acemel RD, Maeso I, Gómez-Skarmeta JL. Topologically associated domains: a successful scaffold for the evolution of gene regulation in animals. *Wiley Interdiscip Rev Dev Biol.* 2017 May;6(3). doi: 10.1002/wdev.265.
7. Proyecto Encode: <https://www.genome.gov/10005107/the-encode-project-encyclopedia-of-dna-elements/>
8. Proyecto Roadmap Epigenomics: <http://www.roadmapepigenomics.org/>
9. Proyecto ModeEncode: <http://www.modencode.org/>
10. Proyecto Mouse Encode: <http://www.mouseencode.org/>
11. Proyecto 4D nucleome: <https://commonfund.nih.gov/4dnucleome>

Figura. Región del genoma de pez cebra mostrando los genes six2a y six3a. De abajo a arriba se muestran los genes de la zona, los sitios donde se localizan las marcas epigenómicas H3K3me3 (promotores activos) y H3K27ac (elementos reguladores activos), las zonas de cromatina abierta (ATAC-seq) y los contactos de los promotores six2a (azul) y six3a (rojo), determinados por técnicas de captura de conformación cromosómica. Los asteriscos azules y naranjas muestran elementos reguladores activos en momentos diferentes del desarrollo embrionario. Nótese que los promotores de los genes six2a y six3a contactan zonas opuestas del genoma, donde se encuentran sus correspondientes elementos reguladores. Estas dos zonas definen dos TADs diferentes.

