



La secuenciación de los genomas de dos tortugas: *Pelodiscus sinensis* y *Chelonia mydas*

Juan Pascual Anaya

Investigador científico contratado. Laboratory for Evolutionary Morphology,
RIKEN Center for Developmental Biology.
Minatojima-minami, 3-2-2, 650-0047
Kobe, Japan
jpascualanaya@gmail.com

139

En el año 2010, justo acabada la tesis doctoral en la Universitat de Barcelona (a la que me mudé tras licenciarme en Biología por la Universidad de Málaga), empecé mi estancia post-doctoral en el Centro de Biología del Desarrollo del RIKEN, en el laboratorio del Dr. Shigeru Kuratani, probablemente el investigador más destacado de la última década en Japón en el campo de la evolución del desarrollo, y uno de los más importantes del mundo en el campo de la embriología comparada.

El laboratorio del Dr. Kuratani estudia los cambios morfológicos que se han producido durante la evolución de los vertebrados, prestando especial atención a aquellas innovaciones que suponen unos cambios drásticos en el bauplan de estos. Por ejemplo, la aparición de la mandíbula, el origen de la cabeza (y al mismo tiempo estructuras clave relacionadas como la cadera anterior y el cuello) y, sobre todo, el origen del caparazón de la tortuga han sido algunos de los ejemplos estudiados en profundidad en su laboratorio.

Como es bien sabido, la evolución consiste en descendencia con modificación, lo cual implica un hecho muy importante en el marco evolutivo: que todo lo nuevo deriva de algo viejo. Sin embargo, esto no está tan claro para ciertas estructuras, las que se denominan innovaciones verdaderas; es decir, que no aparecen claramente a partir de la modificación de una estructura que estaba presente anteriormente. Una de estas estructuras es el caparazón de la tortuga.

En el año 2005, el laboratorio del Dr. Kuratani publicó un trabajo sobre genes que se expresan específicamente en la cresta ectodérmica del caparazón (CR, del inglés *carapacial ridge*), la estructura embrionaria que derivará en el caparazón, descubriendo la implicación de genes clave como *LEF1*, un factor de transcripción que se activa al final de la vía de señalización de WNT [1]. La expresión de estos genes, junto con la similitud histológica entre la CR y la cresta apical ectodérmica de los primordios de las extremidades llevaron a pensar que el origen de la CR se encuentra asociado a la co-opción de redes génicas ya establecidas en el desarrollo de las extremidades [2]. Sin embargo, este trabajo estaba limitado a la poca cantidad de información genética que se tenía entonces sobre las tortugas. Y este fue el principal motivo que llevó al grupo a plantearse un proyecto de secuenciación de genoma, bajo la coordinación del Dr. Naoki Irie, de la tortuga china de caparazón blando (*Pelodiscus sinensis*), modelo con el que trabajamos. El trabajo ha sido recientemente publicado en *Nature Genetics*, y es de acceso libre [3], y en los siguientes párrafos explicaré brevemente la historia del proyecto, los resultados más importantes y, sobre todo, mi participación científica en él.

En el año 2010 el proyecto empezó mediante la concesión de financiación para su realización por parte del gobierno japonés. Tras varios contactos con algunos proveedores, llegamos a la conclusión de que aunque las nuevas técnicas de secuenciación han abaratado mucho los proyectos de genómica, el presupuesto no llegaba para hacer todo lo que queríamos solos. Así pues, en enero de 2011 fuimos a la XIX conferencia internacional "*Plant and Animal Genome*" en San Diego, California (EEUU). Y fue allí donde entablamos los contactos para asociarnos al *Beijing Genome Institute* (BGI) de China y al *Instituto Europeo de Bioinformática* (EBI, de sus siglas en inglés). El enorme esfuerzo que supone la secuenciación de un genoma *de novo* (no la secuenciación de genomas de individuos para los que existe ya un genoma de referencia, como el caso del ser humano), no sólo económico, sino de recursos humanos, hizo que el número de instituciones participantes aumentara durante los siguientes seis meses, englobando finalmente a nueve centros de investigación en total.

Además, dentro del BGI ya había otro grupo que había empezado a secuenciar el genoma de la tortuga verde marina (*Chelonia mydas*) y decidimos unir esfuerzos con el objetivo de presentar un trabajo más completo con dos genomas que diera ideas sobre el grupo de las tortugas en general y no solo sobre una especie en particular. Queda de manifiesto por tanto la importancia de establecer fructíferas colaboraciones para la realización de proyectos de esta envergadura.

Así pues, el trabajo a realizar quedó repartido principalmente entre el RIKEN, el BGI y el EBI de la siguiente manera: el BGI se encargaría de la secuenciación y ensamblaje del genoma, así como de la realización de una anotación (descripción estructural del genoma: coordenadas génicas de los diferentes elementos génicos: genes, elementos repetitivos, RNAs no codificantes, etc...) preliminar; el EBI se encargaría de la realización de una página *web* donde se daría un soporte gráfico para análisis así como de una anotación más profunda utilizando *RNA-seq* (secuencias de RNA producidas mediante técnicas de última generación de secuenciación a partir del RNA mensajero que producen millones de "lecturas" que luego se alinean frente al genoma para establecer así el origen genómico de los transcritos); y en el RIKEN nos encargaríamos de un análisis funcional del genoma.

Entre los diferentes análisis a realizar, el más importante era aprovechar la ingente cantidad de información que puede proporcionar un genoma para establecer definitivamente la posición filogenética de las tortugas, posición que siempre ha sido dudosa usando tanto datos morfológicos como moleculares. Las tortugas han sido clasificadas como (i) una rama temprana de los reptiles (anápsidos) colocándolos a la base de los diápsidos, (ii) como grupo hermano de arcosaurios (cocodrilos y aves) o (iii) como grupo hermano de lepidosaurios (lagartos, serpientes y tuataras), y no sólo basándose en aspectos morfológicos, sino también moleculares (genes que codifican para proteínas, microRNAs o incluso elementos no codificantes conservados). En nuestro caso usamos la secuencia encadenada de 1113 genes de las dos tortugas estudiadas así como de otros diez genomas de vertebrados secuenciados (pollo, diamante mandarín, cocodrilo marino, aligátor americano, lagarto anoles, perro, humano, ornitorrinco, la rana *Xenopus tropicalis* y el medaka común). Los diferentes análisis filogenéticos usando este set de datos, el mayor hasta la fecha, emplazan inequívocamente a las tortugas como grupo hermano de arcosaurios (Figura 1). El tiempo de divergencia estimado para las tortugas, ajustado mediante el uso de fósiles, es de aproximadamente 260 millones de año. Este periodo coincide con el final del Pérmico y el inicio del Triásico, momento que en el que se produce una gran extinción y que podría estar asociado con la aparición y éxito posterior de las tortugas.

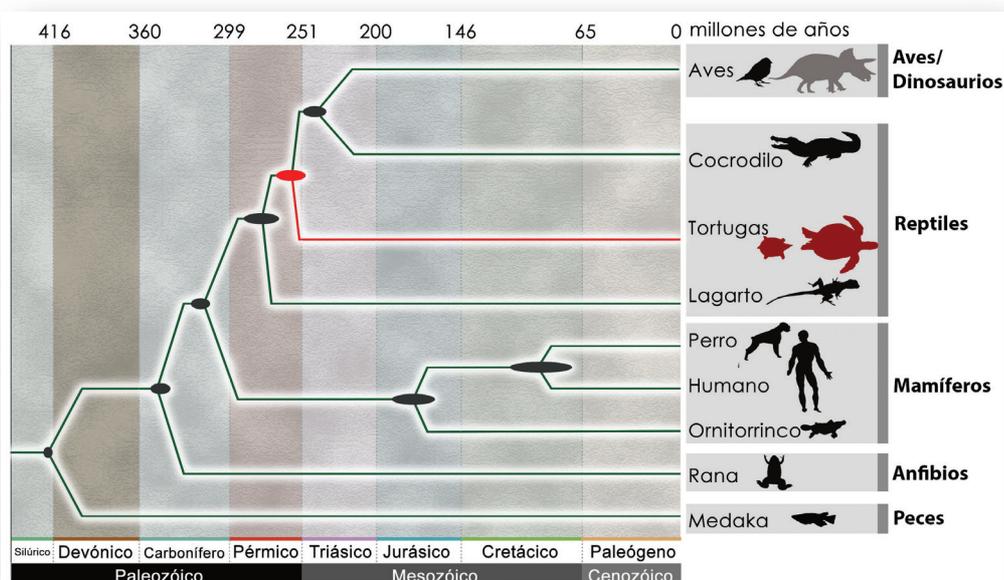


Figura 1: Posición filogenética y tiempo de divergencia de las tortugas respecto a otros vertebrados, usando la secuencia encadenada de 1113 genes. Imagen cedida por el Dr. Naoki Irie, coordinador del proyecto.

Como objetivos del proyecto se encontraban además identificar rasgos genómicos específicos de las tortugas, por un lado, e intentar aprovechar la excelente herramienta que supone un genoma para identificar elementos moleculares del desarrollo embrionario específicos para la tortuga. Respecto al primer objetivo, es reseñable la expansión de genes que codifican para receptores de odorantes que ha ocurrido en tortugas, especialmente en *P. sinensis* que llega hasta niveles similares o incluso superiores a los de mamíferos. Esta expansión es específica e independiente de tortugas (sobre todo de receptores tipo α) y queda mejor reflejado en el gráfico de la figura 2.

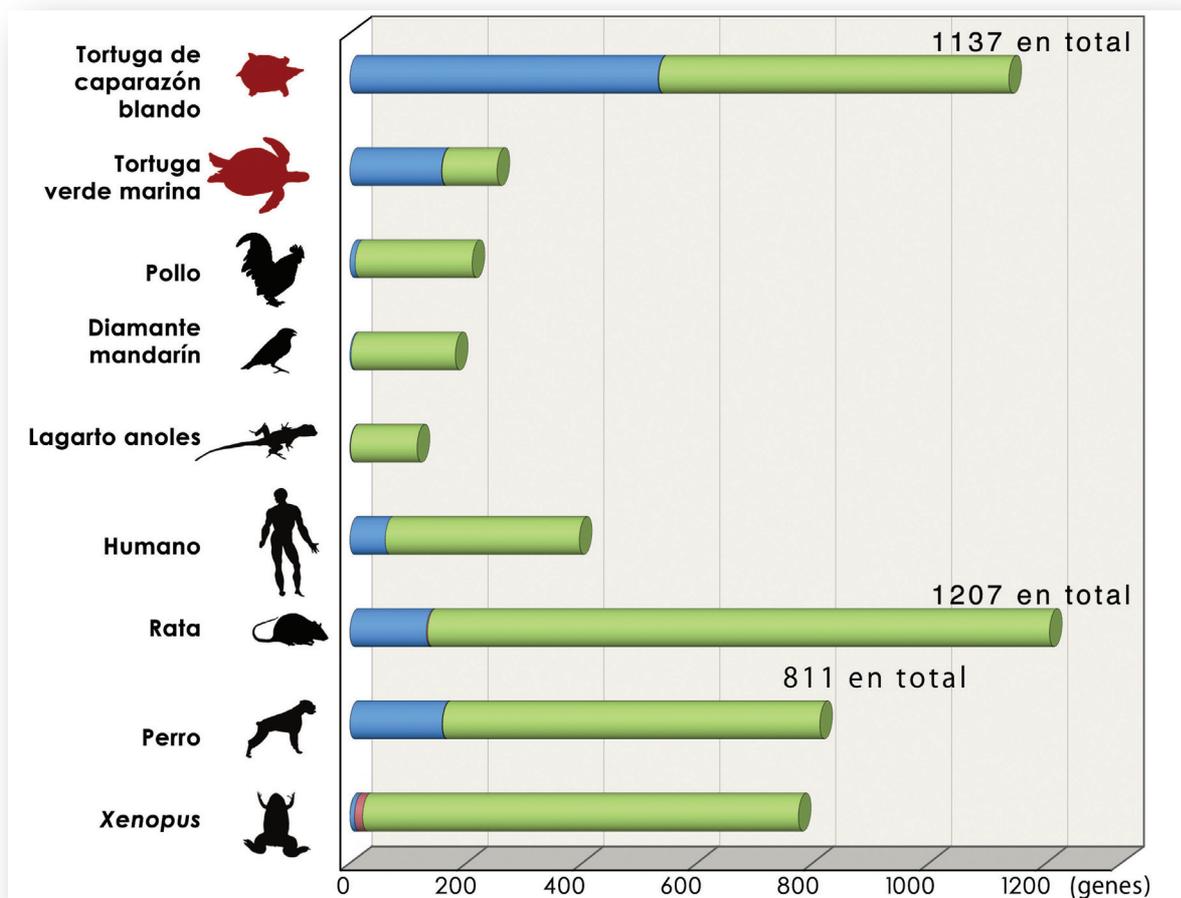


Figura 2: Gráfico que muestra la expansión de los genes que codifican para los receptores de odorantes en diferentes genomas de vertebrados. En azul, receptores tipo α ; en verde, receptores tipo γ ; en rojo, receptores tipo β . Imagen cedida por el Dr. Naaki Irie, coordinador del proyecto

En el segundo objetivo es donde se enmarcó mi función dentro del proyecto. Mi participación se basó en identificar los aspectos moleculares relacionados con el desarrollo del caparazón. Como he explicado antes, en un trabajo anterior se había identificado un posible papel de la vía de señalización WNT en el desarrollo del caparazón, pero sin embargo no se había identificado hasta la fecha la expresión asociada de ningún ligando WNT. Así pues, decidí buscar todos los genes que codifican para ligandos WNT en el genoma de las tortugas, y analizar su expresión durante el desarrollo. En total pude anotar 20 genes que codifican para WNT, incluyendo *WNT10B*, que parece haberse perdido en el linaje de las aves, y *WNT11B*, perdido en mamíferos. Tras analizar la expresión de los 20 genes, observamos que *WNT5A* era el único expresado en la CR. Sorprendentemente, *WNT5A* también se encuentra activo en el desarrollo de las extremidades, lo que da fuerza a las hipótesis anteriores que sugieren una co-optimación de las redes génicas de las extremidades a la CR. Por otro lado, decidí también investigar el papel de un tipo de moléculas cuyo rol en el desarrollo embrionario es aún bastante desconocido: los microRNAs. Los microRNAs son unas moléculas de RNA pequeñas, de entre 21-23 nucleótidos que regulan la traducción de los genes a proteínas mediante mecanismos de represión en los que no entraré en detalle. Se separó la fase de RNAs pequeños de una extracción de RNAs de tres tejidos diferentes (la CR, las extremidades y la pared lateral del cuerpo) de embriones en el estadio en el que la CR empieza a desarrollarse. Luego, esta fase de RNAs pequeños fue secuenciada masivamente por técnicas de *small RNA-seq* (igual al *RNA-seq*, pero con las moléculas pequeñas), y por último, mediante técnicas de bioinformática pude predecir, usando los resultados del *small RNA-seq* y el genoma, microRNAs que posiblemente están expresados en estos tejidos. Finalmente, comparando la presencia o ausencia de microRNAs entre los tres tejidos, se pueden identificar microRNAs específicamente detectados en un tejido y no en los demás. En total, pude identificar la presencia de 212 microRNAs específicos de la CR, y que podrían tener algún papel en su desarrollo.

En conclusión, lo que es más valioso de este trabajo es la obtención de unas herramientas tan importantes como los genomas de dos tortugas: una muy utilizada en estudios de desarrollo, *P. sinensis*, y la otra que es primordial en programas de conservación de la biosfera, como es la tortuga verde marina, *C. mydas*, en peligro de extinción. Aunque en nuestro trabajo ya presentamos algunos resultados sorprendentes y que podrían estar relacionados con algunos de los rasgos más llamativos de las tortugas como el caparazón, estoy seguro que los resultados más importantes en el campo están aún por llegar y en ellos la utilización de estos genomas jugarán un papel fundamental.



Bibliografía citada:

1. Kuraku S, Usuda R, & Kuratani S (2005) Comprehensive survey of carapacial ridge-specific genes in turtle implies co-option of some regulatory genes in carapace evolution. *Evolution & Development* 7(1):3-17.
2. Burke A (1989) Development of the turtle carapace: implications for the evolution of a novel bauplan. *J Morphol* 199:363-378.
3. Wang Z, et al. (2013) The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nature Genetics*.