

## Club de revistas

### Reseña sobre los artículos

Foley, N. M., Mason, V. C., Harris, A. J., Bredemeyer, K. R., Damas, J., Lewin, H. A., Eizirik, E., Gatesy, J., Karlsson, E. K., Lindblad-Toh, K., Zoonomia Consortium., Springer, M. S. Murphy, W. J. (2023). A genomic timescale for placental mammal evolution. *Science*, 380(6643), eabl8189. <https://doi.org/10.1126/science.abl8189>

Zoonomia Consortium (2020). A comparative genomics multitool for scientific discovery and conservation. *Nature*, 587(7833), 240-245. <https://doi.org/10.1038/s41586-020-2876-6>



## Proyecto Zoonomia: los genomas de cientos de mamíferos dan pistas claras sobre su evolución

### Zoonomia Project: Genomes of hundreds of mammals give clear clues about their evolution

Los mamíferos placentados incluyen más de 6.000 especies vivientes con una amplia diversidad morfológica y ecológica que incluye desde murciélagos de dos gramos (*Craseonycteris thonglongyai*) hasta ballenas de más de 150 toneladas (*Balaenoptera musculus*). En los genomas de los organismos no sólo se encuentran los rasgos que permiten entender su funcionamiento, sino también pistas sobre su evolución. Mediante la genómica comparativa de mamíferos es posible conocer cómo los genomas de este grupo evolucionaron y descubrir las similitudes y diferencias que tenemos nosotros los *Homo sapiens* con respecto a las demás especies de nuestro clado. El Proyecto Zoonomia (Zoonomia, 2023) ha puesto a disposición de la comunidad científica los datos y herramientas para comparar genomas de mamíferos placentados, con el fin de explorar aspectos de su biología y descubrir variantes genómicas que se relacionen con cambios fenotípicos, enfocándose especialmente en identificar aquellos relacionados con riesgos de enfermedad en nuestra especie (Zoonomia Consortium, 2020).

El Proyecto Zoonomia provee ensamblajes nuevos del genoma de más de 130 especies de mamíferos placentados y el alineamiento de 241 especies en una matriz de datos con información del ADN de cada especie en relación con todas las demás, incluidos representantes del 82 % de las familias de mamíferos placentados (Zoonomia Consortium, 2020). En este proyecto participan más de 30 instituciones, 28 de las cuales colaboraron en la obtención de las muestras biológicas, siendo central el rol del zoológico de San Diego en California, Estados Unidos, el cual mantiene desde hace varias décadas germoplasma de varias especies de animales (Zoonomia Consortium, 2020).

Los genomas de las especies del proyecto se obtuvieron mediante la tecnología de secuenciación Ilumina con adaptadores que no requieren de PCR y que usan ADN de calidad media (es decir, muestras de ADN de fragmentos aislados que tienen más de 5.000 pares de bases). A pesar de que los requerimientos de calidad de las muestras de ADN no eran muy altos, el 20 % de las muestras preseleccionadas no pudieron emplearse, pues el

ADN no alcanzaba esa calidad (**Zoonomia Consortium**, 2020). El ADN de las muestras que cumplía con la calidad se fragmentó usando el sistema Covaris E220; luego se determinó la distribución de los tamaños de los fragmentos con un sistema de electroforesis automatizada de alta resolución (Agilent 2100 Bioanalyzer) para construir las librerías que después se secuenciaron en el sistema Illumina HiSeq 2500.

Para ensamblar el genoma de cada especie del proyecto, se implementó el procedimiento bioinformático DISCOVAR *de novo* a partir de la información obtenida de la secuenciación en el sistema Illumina (**Weisenfeld et al.**, 2014). El promedio de la profundidad de cobertura de todos los genomas obtenidos fue de 40X (DE+/- 14X). Esta cobertura indica que cada nucleótido de los genomas se leyó alrededor de 40 veces durante el procedimiento de secuenciación, permitiendo, así, diferenciar entre variaciones reales y errores de secuenciación. Los genomas ensamblados se encuentran en <https://zoonomiaproject.org/the-mammal-tree-list-view/>, e incluyen varias especies neotropicales como el jaguar (*Panthera onca*), la danta (*Tapirus terrestris*), el delfín rosado (*Inia geoffrensis*), el guatín (*Dasyprocta punctata*), el chigüiro (*Hydrochoerus hydrochaeris*), murciélagos (*Noctilio leporinus* y *Carollia perspicillata*), monos (*Ateles geoffroyi* y *Alouatta palliata*), osos perezosos (*Bradypus variegatus* y *Choloepus didactylus*) y osos hormigueros (*Myrmecophaga tridactyla* y *Tamandua tetradactyla*).

El alineamiento de los genomas de las 241 especies, incluidos los obtenidos por fuera del proyecto, se encuentra disponible en <https://cglgenomics.ucsc.edu/data/cactus/>. Dada la escala de los datos y de la diversidad taxonómica, construir este alineamiento fue un proceso complejo que requirió un procedimiento particular (**Armstrong et al.**, 2020, **Zoonomia Consortium**, 2020). En abril del 2023 la revista Science publicó una sección especial con casi una docena de artículos que presentaban los descubrimientos hechos a partir de los datos del Proyecto Zoonomia y que se encuentran reseñados en las perspectivas de **Upham & Landis** (2023) y **Romero** (2023). Comento a continuación un artículo en el que se usa la información genómica para reconstruir la historia evolutiva de los mamíferos placentados (**Foley et al.**, 2023).

El estudio filogenético de **Foley et al.** (2023) se basa en el alineamiento del proyecto Zoonomia, pero los autores, además, generaron una serie adicional de alineamientos. Por ejemplo, los de sitios neutrales o casi neutrales (es decir, que no codifican a proteína), empleando tres genomas de referencia (de humano, de perro y de un “ancestro hipotético”), así como alineamientos con un número menor de taxones para minimizar los datos ausentes en el alineamiento original. También hicieron alineamientos de cromosomas diferentes para analizar posibles incongruencias entre las filogenias derivadas, cuyo objeto era conocer qué tan robustos podían ser los resultados al emplear distintos esquemas de datos, o inferir los procesos evolutivos ocurridos que generarían incongruencias en los resultados obtenidos con diferentes tipos de datos. Además del análisis de la información de las secuencias genómicas, también se analizaron dos tipos de variantes estructurales del genoma (deleciones de más de 10 pares de bases y rearrreglos cromosómicos), las cuales evolucionan más lentamente que las sustituciones de los nucleótidos y fueron una fuente independiente de información para corroborar los resultados filogenéticos basados en los alineamientos. Los tiempos de divergencia en la filogenia se calcularon en el programa PALM (**Yang**, 2007), considerando las edades de 37 fósiles y otros aspectos que pueden generar variaciones en las edades o en la adscripción filogenética de los taxones fósiles.

Los árboles filogenéticos resultantes de los diferentes alineamientos fueron similares, e indicaban que el origen de los mamíferos placentados se dio hace 102 millones de años (intervalo de confianza: 114,5 a 90,4 millones de años). Sin embargo, se presentaron casos con cambios en la topología en diferentes análisis, particularmente dentro del clado Afrotheria, que incluye a los manatíes (*Trichechus*) distribuidos en el Mar Caribe y en algunos ríos del Neotrópico. En los análisis de los alineamientos de diferentes cromosomas los autores compararon los resultados del cromosoma X y de tres autosomas y encontraron que los clados mayores, que incluyen varios órdenes de mamíferos (clados superordinales),

mostraban muy poca discordancia en relación con el cromosoma analizado. Esta poca discordancia en esos clados superordinales (que se diversificaron en el Cretáceo tardío hace 100 a 66 millones de años), puede ser una señal de procesos de especiación alopátrica que aislaron pequeñas poblaciones de los ancestros de mamíferos placentados en diferentes fragmentos de las masas continentales de Gondwana y Laurasia y llevó al patrón biogeográfico de distribución actual de tales clados (Afrotheria en África, Xenarthra en Suramérica, Laurasiatheria y Euarchontoglires en Eurasia y Norte América).

La divergencia inicial de los cuatro clados mayores coincide en el tiempo con la separación de las masas continentales y el aumento en el nivel de los mares, es decir, el Cretáceo tardío, cuando se subdividieron las superficies de tierra emergidas. Estos resultados contrastan con otros modelos previos de diversificación de mamíferos, pero se sostienen frente a las variaciones en los datos y las calibraciones de las edades de los fósiles. Además, los datos usados en estos análisis incluyen más de 40.000 sitios informativos, o sea, diez veces más de lo requerido en el modelamiento de datos para recuperar tiempos de divergencia precisos de un número comparable de taxones.

Un segundo pulso de diversificación al interior de los clados, correspondiente al nivel taxonómico de orden, se debió haber dado inmediatamente después del evento de extinción ocurrido en la transición Cretáceo-Paleógeno hace 65 a 53 millones de años (famoso por la extinción de los dinosaurios) y que también coincide con la emergencia rápida de masas continentales. En contraste con la diversificación inicial de los mamíferos placentados, que muestra una clara división de los linajes superordinales, la diversificación posterior durante el Cenozoico (últimos 65 millones de años), muestra altos niveles de discordancia filogenética. Por ejemplo, la radiación de los primates evidencia el conflicto topológico entre autosomas y el cromosoma X, algo que es característico cuando hay sorteo incompleto de linajes. Este sorteo de linajes es común cuando hay especies ancestrales con poblaciones muy grandes que se subdividen varias veces en una sucesión rápida. Esto lleva a que en el genoma de varias especies descendientes haya parte de la variación genómica de la especie ancestral, pero la similitud de los genomas de esas especies descendientes no necesariamente se debe a una relación histórica cercana. Es decir, un proceso así lleva a que haya incongruencias en la señal filogenética de diferentes cromosomas o regiones del genoma. Otro proceso evolutivo que también causa incongruencias es la especiación con flujo génico o introgresión entre especies, lo que parece haber ocurrido en el orden de los roedores. Los autores proponen que esta discordancia entre regiones en los genomas se debe a que las especies ancestrales tenían grandes tamaños poblacionales y distribuciones geográficas muy amplias, lo que habría favorecido los procesos mencionados.

Aunque el muestreo del proyecto Zoonomia es taxonómicamente amplio e incluye varias especies neotropicales, la colaboración con instituciones latinoamericanas será necesaria para ampliar la cobertura de especies. De hecho, Brasil y Alemania lideran otro consorcio para estudiar la genómica adaptativa de especies neotropicales que no se limita solo a mamíferos placentados (**GenoTropics**, 2023), en tanto que Colombia se ha vinculado al consorcio del Proyecto del Biogenoma de la Tierra (**Huddart et al.**, 2022), lo que demuestra cómo la genómica empieza a hacerse presente en las agendas científicas de varios países. Por último, vale la pena anotar que en junio del 2023, la misma revista Science publicó otra serie especial de artículos sobre genómica de primates que incluye los análisis de genomas de más de 200 especies, lo cual evidencia la rapidez con que se están generando datos de especies relacionadas con nosotros. Este gran volumen de información permite un mejor entendimiento de la diversidad de los mamíferos, con lo que se vislumbra, asimismo, un sinfín de aplicaciones beneficiosas para los humanos.

### **Enrique Arbeláez-Cortés**

Grupo de Estudios en Biodiversidad, Escuela de Biología, Universidad Industrial de Santander, Bucaramanga, Colombia

## Agradecimientos

A Martha Patricia Ramírez Pinilla por la amable invitación a presentar esta reseña.

## Referencias

- Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I. T., Novak, A. M., Deran, A., Fang, Q Xie, D., Feng, S., Stiller, J., Genereux, D., Johnson, J., Marinescu, V., Alföldi, J., Harris, R., Lindblad-Toh, K., Haussler, D., Karlsson, E., Jarvis, E., Zhang, G., Paten, B.** (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. *Nature*, 587(7833), 246-251.
- Foley, N. M., Mason, V. C., Harris, A. J., Bredemeyer, K. R., Damas, J., Lewin, H. A., Eizirik, E., Gatesy, J., Karlsson, E. K., Lindblad-Toh, K., Zoonomia Consortium., Springer, M. S. Murphy, W. J.** (2023). A genomic timescale for placental mammal evolution. *Science*, 380(6643), eabl8189. <https://doi.org/10.1126/science.abl8189>
- Zoonomia Consortium** (2020). A comparative genomics multitool for scientific discovery and conservation. *Nature*, 587(7833), 240-245. <https://doi.org/10.1038/s41586-020-2876-6>
- GenoTropics** (2023). GenoTropics: Adaptive genomics for neotropical species. <https://www.genotropics.org/>
- Huddart, J. E., Crawford, A. J., Luna-Tapia, A. L., Restrepo, S., Di Palma, F.** (2022). EBP-Colombia and the bioeconomy: Genomics in the service of biodiversity conservation and sustainable development. *Proceedings of the National Academy of Sciences*, 119(4), e2115641119.
- Romero, I. G.** (2023). Seeing humans through an evolutionary lens. *Science*, 380 (6643), 360-361.
- Upham, N. S., Landis, M. J.** (2023). Genomics expands the mammalverse. *Science*, 380(6643), 358-359.
- Yang, Z.** (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8), 1586-1591.
- Zoonomia** (2023). Zoonomia Project. <https://zoonomiaproject.org/the-project/>