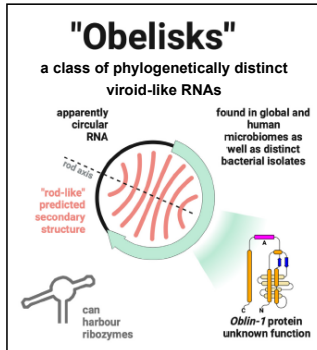


Club de revistas/Journal Club

Graphical abstract



Reseña sobre el artículo

Zheludev, I.N., Edgar, R.C., Lopez-Galiano, M.J., de la Peña, M., Babaian, A., Bhatt, A.S., Fire, A.Z. (2024). Viroid-like colonists of human microbiomes. *Cell*, 187(23),6521-6536.e18.

Explorando la materia oscura microbiana

Exploring microbial dark matter

El mundo microbiano, invisible al ojo, es amplio y misterioso. Parte de él está compuesto por partículas virales cuyo material genético puede ser ADN o ARN (Koonin *et al.*, 2021). Los virus son ubicuos y abundantes y participan en la transferencia de genes entre organismos, por lo cual son importantes en su evolución. Las primeras descripciones a finales del siglo 19 definieron los virus como partículas infecciosas minúsculas capaces de atravesar los filtros que retienen las células de mayor tamaño. Luego se describieron como parásitos intracelulares dependientes de una célula hospedera para su crecimiento y replicación. Hoy se sabe que el mundo de los virus, la virósfera, es extenso y diverso (Edgar *et al.*, 2022; Zayed *et al.*, 2022) y, en muchos casos, desafía las descripciones iniciales (Koonin *et al.*, 2021).

A los virus se le suman otros elementos microscópicos como las moléculas de ADN o ARN satelital y los viroides. Estos últimos consisten en ARN cortos de entre 220 y 450 nucleótidos y, a diferencia de los virus, no codifican para proteínas, pues su replicación depende de la ARN polimerasa de la célula hospedera. En un trabajo reciente, Zheludev y colaboradores, de la Universidad de Stanford en Estados Unidos, evidenciaron que el conjunto de elementos microscópicos es todavía más extenso y desconocido de lo esperado (Zheludev *et al.*, 2024). Utilizando herramientas bioinformáticas y bases de datos sólidas y abiertas, identificaron un nuevo grupo de elementos microscópicos cuyo material genómico es el ARN, al cual llamaron obeliscos (*obelisks*).

El equipo de investigadores se propuso identificar posibles viroides y agentes similares cuyos pequeños genomas contienen poca información. En comparación con los virus de ARN (Riboviria), se conocen pocas partículas del tipo viroides debido, en gran medida, a que son difíciles de identificar. La ausencia de moléculas esenciales para su replicación, como las ARN polimerasas, hace inviable su identificación mediante búsquedas basadas en la homología. Por lo tanto, los investigadores procedieron a diseñar una nueva aproximación bioinformática independiente de la homología a la que llamaron *Viroid Nominator* o *VNom*, capaz de identificar secuencias con características deseadas, como genomas pequeños e formación de secuencias circulares, a partir de datos de secuenciación del ARN (RNA-seq).

La búsqueda inicial de elementos del tipo viroides se hizo utilizando *VNom* para analizar los metatranscriptomas, o de ARN metagenómico, proveniente de muestras longitudinales de heces de 104 donantes humanos. Con este análisis encontraron 15 secuencias novedosas, es decir, sin homología con secuencias presentes en las bases de datos. Estos ARN estaban presentes en siete de las 104 muestras analizadas; tenían aproximadamente 1.000 nucleótidos (nt) de longitud y una estructura secundaria predicha en forma de vara extendida (*rod*) debido al apareamiento de bases, motivo por el que se les dio el nombre de *obelisk-alpha* (*obelisk- α*). El ARN del *obelisk- α* codifica para dos proteínas, *oblina-1* y *oblina-2*, de 202 y 53 aminoácidos, respectivamente, que tampoco tenían secuencias homólogas en las bases de datos de referencia. Estas características sugieren que el *obelisk- α* es un tipo de viroide novedoso presente en microbiomas humanos.

Pero ¿en dónde más podrían encontrarse este tipo de elementos? Se hizo entonces una búsqueda más extensa utilizando k-meros (subsecuencias de longitud k) para indagar 3,2 millones de secuencias metagenómicas de la base de datos SRA (Sequence Read Archive,

el repositorio público más grande de secuenciación de alto rendimiento). De esta manera se identificaron 21 ARN adicionales de características similares, incluido un elemento con leves diferencias que denominaron *obelisk beta* (*obelisk-β*). Estos elementos se encontraron únicamente en bases de datos de ARN, y no de ADN, lo cual confirma que sus genomas son de ARN, al igual que los viriones. Ambos elementos, el *obelisk-α* y el *obelisk-β*, tienen genomas pequeños que forman estructuras de vara y contienen las proteínas *obli*-1 y -2. Al buscar homólogos de estas dos proteínas en otros repositorios, en este caso de ensamblajes de virus de ARN (*RNA deep virome assemblage*, RDVA), identificaron 38.500 posibles obeliscos. Con esta información ampliada volvieron a sondear la base de datos SRA usando modelos de Markov (pHMM) y encontraron obeliscos con una o ambas de las proteínas *obli*-1 y *obli*-2. Por último, hicieron una búsqueda para identificar ribozimas, moléculas de ARN que actúan como enzimas y participan en el ciclo de replicación de los viroides, y encontraron 339 obeliscos con ribozimas. Así pues, estos elementos codifican para una o dos proteínas y, en algunos casos, para ribozimas. Como resultado de las múltiples búsquedas bioinformáticas utilizando datos de diversos ambientes y geografías, el grupo de investigadores identificó 29.959 obeliscos únicos presentes en diferentes nichos ecológicos además del microbioma humano. En resumen, estos elementos no tienen homología con virus u otros elementos descubiertos anteriormente, forman un grupo filogenético propio, son diversos y de distribución global.

Con el fin de determinar características distintivas en las recién identificadas proteínas *obli* 1 y *obli* 2 codificadas por los obeliscos, las dos se sometieron a predicciones estructurales utilizando ColabFold (Mirdita *et al.*, 2022), una implementación de AlphaFold (Jumper *et al.*, 2021), que es un desarrollo basado en inteligencia artificial cuya capacidad para predecir estructuras proteicas fue merecedora del Premio Nobel de Química en 2024. A pesar de sus variaciones, las diferentes *obli* 1 contienen una región conservada que carece de estructura terciaria evidente y otro dominio que forma una estructura globular. Por su parte, las *obli* 2 poseen una α -hélice con un motivo llamado cremallera de leucina (*leucine zipper*) que sirve para reunir y generar oligomerización de proteínas. Los autores especulan que, en conjunto con las ribozimas presentes en algunos obeliscos, dichas proteínas podrían participar en la replicación.

Si bien los análisis bioinformáticos revelan que estos elementos están presentes en diversas muestras depositadas en bases de datos, es importante identificar células que los alberguen y, si posible, verificar las predicciones bioinformáticas en el laboratorio. Con este fin, los investigadores se centraron en microbiomas menos complejos e identificaron una bacteria comensal del microbioma oral, *Streptococcus sanguinis*, que contenía un elemento con características de obelisco (tamaño, estructura secundaria, homología con las *obli*-1 y complementariedad de secuencia). Utilizando RT-PCR lograron amplificar la secuencia del obelisco, verificando así su presencia en un hospedero bacteriano. También encontraron que algunos cultivos perdían el obelisco durante el crecimiento. Sin embargo, la falta del obelisco no alteró el crecimiento de *S. sanguinis*, lo que sugiere que no es esencial para su hospedero. El hallazgo de un sistema hospedero-obelisco permite realizar estudios en el laboratorio para entender aspectos como los mecanismos de replicación y su posible transmisión entre células.

Quedan muchas preguntas por resolver sobre estos obeliscos. Por ejemplo, ¿cuáles son los hospederos en ambientes naturales? ¿Cómo interactúan con diversas formas de vida? ¿Cómo se replican y transmiten entre organismos y qué funciones cumplen? Dada la dificultad para detectar elementos novedosos, es posible que este estudio subestime la diversidad real de los obeliscos. También es probable que existan otros tipos de elementos de ARN aún por descubrir. Lo que sí es indudable es que la creatividad para desarrollar estrategias bioinformáticas aunada a la exploración sistemática de la información depositada en bases de datos abiertas, hace posible descifrar la información genómica y

revelar nuevos componentes de nuestra biósfera (**Hou et al.**, 2024). El estudio de Zheludev y colaboradores revela la existencia de elementos invisibles al ojo cuya simple existencia implica procesos evolutivos e interacciones que aún desconocemos, pero que pueden ser relevantes en diversos procesos y formas de vida en nuestro planeta.

María Mercedes Zambrano, Ph. D.

Directora Científica, Corporación Corpogen, Bogotá, Colombia
Dirección de Investigación y Transferencia de Conocimiento, Universidad Central,
Bogotá, Colombia

Referencias

- Edgar, R.C., Taylor, B., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., Banfield, J.F., de la Peña, M., Korobeynikov, A., Chikhi, R., Babaian, A. (2022). Petabase-scale sequence alignment catalyses viral discovery. *Nature*, 602(7895),142-147.
- Gou, Q.Y., Xin, G.Y., Le, S.J., Xia, Y.Y., Zhou, Y.L., Hui, F.M., Pan, Y.F., Eden, J.S., Yang, Z.H., Han, C., Shu, Y.L...Shi, M. (2024). Using artificial intelligence to document the hidden RNA virosphere. *Cell*, 187, 1-14
- Hou, X., He, Y., Fang, P., Mei, S.Q., Xu, Z., Wu, W.C., Tian, J.H., Zhang, S., Zeng, Z.Y., Gou, Q.Y., Xin, G.Y., Le, S.J., Xia, Y.Y., Zhou, Y.L., Hui, F.M., Pan, Y.F., Eden, J.S., Yang, Z.H., Han, C., Shu, Y.L...Shi, M. (2024). Using artificial intelligence to document the hidden RNA virosphere. *Cell*, 187, 1-14
- Koonin, E.V., Dolja, V.V., Krupovic, M., Kuhn, J.H. (2021). Viruses defined by the position of the virosphere within the replicator space. *Microbiology Molecular Biological Review*, 85, e00193-20.
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M. (2022) ColabFold: making protein folding accessible to all. *Nature Methods*, 19(6), 679-682.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589.
- Zayed, A.A., Wainaina, J.M., Domínguez-Huerta, G., Pelletier, E., Guo, J., Mohssen, M., Tian, F., Pratama, A.A., Bolduc, B., Zablocki, O., Carradec, Q., da Silva, C., Labadie, K., Poulain, J., Ruscheweyh, H.J., Salazar, G., Shatoff, E., Coordinators, T.O., Bundschuh, R., Fredrick, K., ..., Sullivan, M.B. (2022). Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome. *Science*, 376(6589),156-162.
- Zheludev, I.N., Edgar, R.C., Lopez-Galiano, M.J., de la Peña, M., Babaian, A., Bhatt, A.S., Fire, A.Z. (2024). Viroid-like colonists of human microbiomes. *Cell*, 187(23),6521-6536.e18.