

## BIOINFORMÁTICA APLICADA – ANÁLISIS DE GENES, PROTEÍNAS Y PANGENOMAS DE BACTERIAS Y EUKARIOTES

Pablo Vinuesa<sup>1</sup>, Javier Rivera<sup>1</sup>, Julio C. Valerdi Negreros<sup>1</sup> y Fulvia Stefany Argueta Zepeda<sup>1</sup>.

<sup>1</sup>Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México (CCG-UNAM), Cuernavaca, Morelos, 62210, México, [vinuesa@ccg.unammx](mailto:vinuesa@ccg.unammx), @pvinmex.

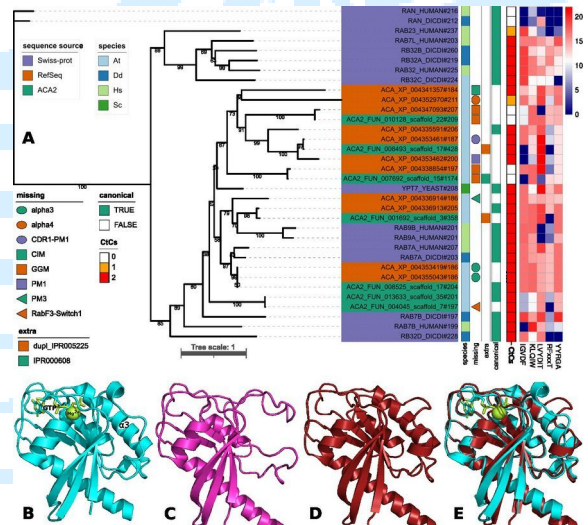
**Palabras clave:** Rab GTPasas, homología profunda, interacción *Stenotrophomonas-Acanthamoeba*

**Introducción.** En años recientes se han publicado una gran cantidad de herramientas bioinformáticas muy sofisticadas, pero de fácil manejo, que son un recurso extraordinario para el análisis de la estructura, función y evolución de genes, proteínas, genomas y pan-genomas. Aquí presentamos unos ejemplos concretos de aplicación de una selección de estas herramientas en un proyecto de investigación biológica enfocado a descubrir y evaluar funcionalmente genes requeridos para la interacción entre *Stenotrophomonas maltophilia* Sm18 y la ameba de vida libre *Acanthamoeba terricola* Neff (antes *A. castellanii*), un fagocito profesional que preda bacterias por fagocitosis. Si bien ambas especies cuentan con genomas completos, su biología está pobremente estudiada, particularmente en lo que respecta a funciones requeridas para la interacción bacteria-hospedero. Los objetivos biológicos del trabajo son: 1) definir el repertorio de las Rab GTPasas (Rabs) y sus efectores, así como complejos de amarre (HOPS) y de fusión (SNAREs) codificados en el genoma de *A. terricola* y potencialmente asociados a la vía fagocítica de la ameba; 2) desarrollar herramientas genéticas para estudiar la interacción y las características fundamentales de la vacuola que aloja a *S. maltophilia* (SmCV) y 3) identificar loci genómicos de la bacteria potencialmente involucrados en la interacción, y analizar funcionalmente una selección de ellos.

**Metodología.** Para la reconstrucción *in silico* de la vía fagocítica de *A. terricola* se buscaron homólogos de proteínas de humano, levadura y *Dictyostelium discoideum* involucradas en ella, haciendo uso de métodos de inferencia de homología profunda basados en modelos ocultos de Markov (HMMs; HMMER3 y HH-suite3) y de alineamientos de estructura-estructura (FoldSeek<sup>1</sup>). Para la reanotación de la estructura de genes de *A. terricola* se usó miniprot, como describimos recientemente<sup>2</sup>. Alineamientos múltiples guiados por estructura

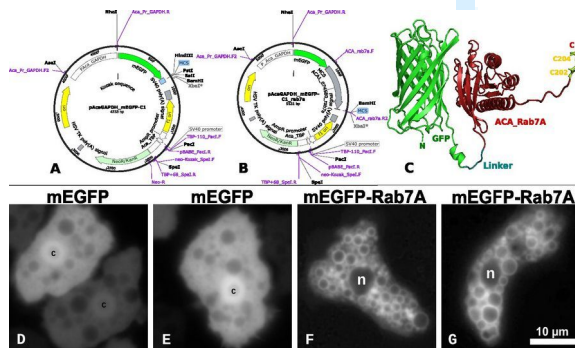
se generaron con mafft-dash<sup>3</sup>. El modelado estructural de proteínas se realizó con AlphaFold2 vía ColabFold<sup>4</sup>. El despliegue de estructuras y su superposición se hizo con PyMOL. Familias de homólogos, pan-genomas y genes específicos del linaje *S. maltophilia* se identificaron con GET\_HOMOLOGUES<sup>5</sup>. Análisis filogenómicos se realizaron con GET\_PHYLOMARKERS<sup>6</sup>.

**Resultados.** Análisis filogenómicos de Rabs identificaron a 11 parálogos de Rab7 en *A. terricola* (Fig 1A), un marcador clave de fagosomas maduros. Análisis estructurales detallados mostraron que 83% de estos genes codifican para Rabs no-cánonicas (Fig. 1A). Demostramos mediante análisis estructural que la anotación de varios de estos genes en el genoma RefSeq de *A. terricola* es incorrecta, corrigiéndola para Rab7A<sup>2</sup> (Fig. 1B-E).



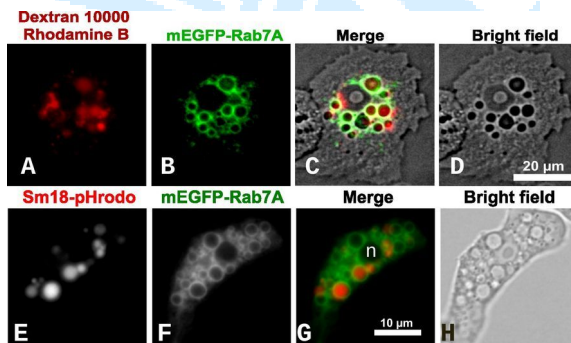
**Fig. 1.** (A) Análisis filogenómico de parálogos de Rab7 de *A. terricola* y organismos de referencia. (B) Estructura de RAB7A\_HUMAN activada con GTP (PDB: 1T91), (C) modelo AF de RAB7A\_ACA, (D) modelado con ColabFold de Rab7A con anotación corregida, y su alineamiento sobre 1T91 (E).

Construimos nuevos plásmidos de expresión para *A. terricola*<sup>2</sup> (Fig. 2A-B) y expresamos una fusión mEGFP-Rab7A en trofozoítos de la ameba (Fig. 2C-G), demostrando que la fusión se localiza en la superficie de vacuolas (Fig 2F-G).



**Fig. 2.** (A-B) Nuevos plásmidos de expresión para *Acanthamoeba*. (C) Modelado estructural con ColabFold de la proteína de fusión mEGFP-Rab7A\_ACA. (D-E) LCI de trofozoítos transfectados con el plásmido vacío (A), o (F-G) que expresan la proteína de fusión mEGFP-Rab7A (C).

Usando trazadores endocíticos y fagocíticos fluorescentes demostramos que la fusión co-localiza con compartimentos endocíticos (Fig. 3A-D) y que *S. maltophilia* marcada *in vivo* con pHrodo se aloja en fagosomas ácidos, Rab7A-positivos<sup>2</sup> (Fig. 2E-H).



**Fig. 3.** (A-D) Co-localización del trazador endocítico dextrano, y (E-H) de células vivas de Sm18 teñidas con pHrodo Red, con vacuolas Rab7A-positivas de *A. terricola*.

Desarrollamos nuevos mini-transposones Tn7 para marcar de manera estable a *S. maltophilia* con mScarlet-l<sup>2</sup>. Con la cepa reportera Sm18::mScarlet-I establecimos un bioensayo de co-cultivo con trofozoítos de *A. terricola* en microplacas de 96 pozos<sup>2</sup>, el cual demostró que Sm18 replica en los trofozoítos,

regresando al medio extracelular por una vía exocítica no lítica<sup>2</sup>.

Resultados preliminares indican que para la vida intracelular Sm18 requiere, entre otros, un sistema de secreción de tipo IV (T4SS) y genes involucrados en homeostasis metálica de Cu y Mn, como revelan bioensayos de co-cultivo de trofozoítos de *A. terricola* con mutantes de *S. maltophilia* en estos loci, generados con nuevos plásmidos que las marcan simultáneamente con proteínas fluorescentes (*en prep.*).

**Conclusiones.** El uso de diversas herramientas bioinformáticas modernas permitió reconstruir *in silico* la vía fagocítica de *A. terricola* con gran detalle, e identificar y corregir la secuencia del ortólogo de Rab7A. Su clonación y expresión como proteína de fusión mEGFP-Rab7A permitió descubrir que *S. maltophilia* establece un nicho de replicación intracelular en vacuolas acidificadas Rab7A-positivas (SmCV), regresando al medio extracelular por una vía exocítica no lítica. *S. maltophilia* es por tanto una bacteria que interactúa extensamente con la vía fagocítica de *A. terricola*, permitiendo el desarrollo de fagosomas maduros. Para ello requiere, entre otros, de un T4SS y genes involucrados en homeostasis metálica.

**Agradecimiento.** Agradecemos el financiamiento recibido del Consejo Nacional de Ciencia y Tecnología (CONACyT México, A1-S-11242) y Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), Universidad Nacional Autónoma de México (DGAPA-PAPIIT, UNAM: IN209321).

#### Bibliografía.

1. van Kempen M, Kim SS, Tumescheit C, et al. 2023. *Nat Biotechnol.* doi:10.1038/s41587-023-01773-0
2. Rivera J, Valerdi-Negreros JC, Vázquez-Enciso DM, Argueta-Zepeda S, Vinuesa P. 2023. *submitted.*
3. Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. 2019. *NAR.* 2019;47(W1):W5-W10. doi:10.1093/NAR/GKZ342
4. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. *Nat Methods.* 2022;19(6):679-682. doi:10.1038/s41592-022-01488-1
5. Contreras-Moreira B, Vinuesa P. 2013. *AEM.* 2013;79(24). doi:10.1128/AEM.02411-13
6. Vinuesa P, Ochoa-Sánchez LE, Contreras-Moreira B. 2018. *Front. Microbiol* 9(MAY). doi:10.3389/fmicb.2018.00771