

## Póster

# Integración de herramientas para la anotación global de genomas procariotas: desde la predicción de genes hasta la definición de rutas metabólicas y reacciones



Carlos Sánchez Casimiro-Soriguer(1), Hikmate Abriouel(2), Nabil Benomar(2) y Antonio J. Pérez-Pulido(1)

(1)Centro Andaluz de Biología del Desarrollo, Universidad Pablo de Olavide-Consejo Superior de Investigaciones Científicas, Junta de Andalucía, Sevilla, Spain.

(2)Área de Microbiología, Departamento de Ciencias de la Salud, Universidad de Jaén. 23071-Jaén.

*Palabras clave:* Anotación, Procariota, Sma3s, Prodigal, Infernal, Pathway Tools

## RESUMEN

**Motivación:** Debido al abaratamiento de la secuenciación de genomas, en los últimos años ha aumentado el número de organismos cuyo genoma ha sido completamente secuenciado, especialmente en el caso de organismos procariotas, con genomas más cortos. Pero una vez que se dispone de la secuencia de ADN de un nuevo genoma, es necesario extraer toda la información que encierran y que puede resolver preguntas biológicas relacionadas con el organismo en estudio, como por ejemplo el tipo de compuestos que pueden metabolizar o la relación evolutiva que presentan con respecto a otros organismos cercanos. Para ello, hay que llevar a cabo una anotación tanto estructural como funcional del genoma, tratando de situar sus elementos genómicos y la función que llevan a cabo, todo lo puede llevar a definir el mapa metabólico del organismo y sus relaciones con la información almacenada en las bases de datos actuales.

**Métodos:** La propuesta de este trabajo para la anotación de un genoma bacteriano sigue el siguiente flujo de trabajo: partiendo del genoma secuenciado de una bacteria se emplea el programa Prodigal para la localización de todos los genes codificantes de proteínas. Para completar el paso de anotación estructural, se localizan también los genes no codificantes de proteínas, utilizando para ello el programa Infernal, entrenado para encontrar genes de ARNr, ARNt y ARNnc. Una vez disponemos de las coordenadas de todos los genes, se pasa a la anotación funcional de aquellos que codifican para proteínas, utilizando para ello el programa Sma3s, el cual ya ha demostrado buenos resultados en genomas bacterianos. Por último, a partir de la anotación generada, se usa el programa Pathway Tools para predecir las rutas metabólicas y reacciones enzimáticas a partir de los "EC number" (códigos de enzimas), los términos GO (Gene Ontology) y la concordancia con los nombres de genes anotados, lo que nos da como resultado final el mapa metabólico completo de la bacteria, así como un mapa genómico que incluye una predicción de operones. Asimismo, las proteínas de membrana son también predichas, mediante una asignación manual de la reacción de transporte.

**Resultados y conclusiones:** El flujo de trabajo propuesto ha sido ensayado ya con dos genomas bacterianos completos, y ha demostrado que en apenas una semana se puede concluir toda la anotación, teniendo una visión global del genoma analizado la cual puede ayudar en análisis más particulares.

## BIBLIOGRAFIA

- Karp, P.D., Paley, S.M., Krummenacker, M., Latendresse, M., Dale, J.M., Lee, T.J., Kaipa, P., Gilham, F., Spaulding, A., Popescu, L., Altman, T., Paulsen, I., Keseler, I.M. & Caspi, R. 2010, "Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology", *Briefings in bioinformatics*, vol. 11, no. 1, pp. 40-79.
- Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W. & Hauser, L.J. 2010, "Prodigal: prokaryotic gene recognition and translation initiation site identification", *BMC bioinformatics*, vol. 11, pp. 119-2105-11-119.
- Munoz-Merida, A., Viguera, E., Claros, M.G., Trelles, O. & Perez-Pulido, A.J. 2014, "Sma3s: a three-step modular annotator for large sequence datasets", *DNA research : an international journal for rapid publication of reports on genes and genomes*, vol. 21, no. 4, pp. 341-353.