

visualizeGO: UNA HERRAMIENTA BIOINFORMÁTICA PARA LA INTERPRETACIÓN Y VISUALIZACIÓN DE TÉRMINOS DE ONTOLOGÍA GÉNICA EN ESTUDIOS ÓMICOS

Alonso-García^{1*}, M., Fonseca^{1,2}, P.A.S., Pelayo¹, R., Arranz¹, J.J., Gutiérrez-Gil¹, B. y Suárez-Vega¹, A.

¹Dpto. Producción Animal, Facultad de Veterinaria, Campus de Vegazana, 24007 León;

²Instituto de Ganadería de Montaña (CSIC-Univ. de León), Finca Marzanas, 24346 Grulleros, León

*maalg@unileon.es

INTRODUCCIÓN

Los análisis de enriquecimiento de ontología génica (GO, del inglés *Gene Ontology*) se utilizan comúnmente para determinar las funciones biológicas de los genes que se obtienen como resultado en los experimentos ómicos. Sin embargo, las listas de términos GO obtenidos en estos análisis pueden ser muy extensas y, en muchas ocasiones, redundantes, lo que complica su interpretación. Por este motivo, hemos desarrollado el paquete de R *visualizeGO* que permite establecer relaciones jerárquicas entre términos GO, disminuir la redundancia y hacer representaciones gráficas comprensibles.

MATERIAL Y MÉTODOS

El paquete *visualizeGO* se ha desarrollado en el lenguaje R (R Core Team, 2024) y la versión estable actual se encuentra disponible en GitHub (<https://github.com/maalg21/visualizeGO>). La función principal del programa es "*clusterGO*". Esta función requiere una lista de términos GO obtenidos a partir de un análisis de enriquecimiento funcional previo. El objetivo de "*clusterGO*" es agrupar términos GO redundantes en clústeres, definidos según el método Silhouette (Rousseeuw, 1987), utilizando la matriz de similitud semántica. A partir de los resultados de *clusterGO*, el paquete ofrece diferentes funciones. Por una parte, la función "*compareGO*" permite comparar clústeres o listas de términos, con el fin de identificar similitudes entre dos condiciones experimentales. Además, la función "*familyGO*", establece la jerarquía completa añadiendo términos parentales e hijos a los términos proporcionados. Esta función es útil para la posterior representación gráfica de los resultados con la función "*visualizeGO*", que facilita la interpretación de los clústeres identificados. Para ilustrar la funcionalidad del paquete, se utilizó como ejemplo una lista de 137 términos GO obtenida a partir del estudio de Alonso-García et al. (2023). En este estudio se realizó un análisis de expresión diferencial entre muestras de tejido adiposo de cordero lechal en dos etapas diferentes de desarrollo. Además, el rendimiento de la herramienta *visualizeGO* se comparó con el paquete *rrvgo* (v.1.18.0) (Sayols, 2023), ampliamente utilizado para la reducción de redundancia en listas de términos enriquecidos.

RESULTADOS Y DISCUSIÓN

El análisis con *visualizeGO* redujo los 137 términos GO obtenidos por Alonso-García et al. (2023) a 54 clústeres de términos agrupados por su similitud semántica. A nivel biológico, los clústeres con más términos están estrechamente vinculados al metabolismo de ácidos grasos y al desarrollo del tejido adiposo. Esto evidencia que, además de reducir la redundancia, el programa resalta procesos clave asociados al estudio utilizado como ejemplo. En términos de rendimiento, al compararlo con *rrvgo*, *visualizeGO* fue más rápido con un tiempo promedio, respectivamente, de $5,50 \pm 0,59$ frente a $2,21 \pm 0,25$ segundos. Además, al comparar los resultados de los dos paquetes mediante la función "*compareGO*" observamos que todos los clústeres detectados por *rrvgo* fueron detectados también por *visualizeGO*. A nivel gráfico, "*visualizeGO*" ofrece la novedad de representar las relaciones jerárquicas entre los términos GO con la agrupación en clústeres incluida en el mismo gráfico.

CONCLUSIÓN

Se presenta *visualizeGO*, una herramienta bioinformática que permite reducir la redundancia de los resultados de los análisis de enriquecimiento funcional facilitando la interpretación biológica de los mismos. Además, ofrece una mayor versatilidad gráfica, incluyendo representaciones jerárquicas y comparativas que no están disponibles en otros paquetes similares.

REFERENCIAS BIBLIOGRÁFICAS

• Alonso-García, M. et al. 2023. Front. Vet. Sci. 10: 1150996. • R Core Team. 2024. R Foundation for Statistical Computing, Vienna (Austria) • Rousseeuw, P.J. 1987. J. Comput. Appl. Math. 20: 53-65 • Sayols, S. 2023. microPublication Biology.

Agradecimientos: Alonso-García, M. está financiada por una beca predoctoral del Fondo Social Europeo Plus, a través del Programa Operativo de Castilla y León, y de la Junta de Castilla y León, mediante la Consejería de Educación.