

IMPACTO DE LA PROFUNDIDAD DE SECUENCIACIÓN EN LA CALIDAD DE LOS DATOS DE ESTUDIOS DE METATAXONOMIA Y ENSAMBLAJE DE METAGENOMAS

Ramayo-Caldas^{1*}, Y., Vourlaki¹, I.T., Clavell-Sansalvador¹, A., Río-López², R., García-Gil³, J., Xifro⁴, X. y Dalmau², A.

¹IRTA, Torre Marimon, Barcelona; ²IRTA, Monells, Girona; ³IDIBGI, Girona; ⁴TargetsLab, UdG, Girona

*yulixaxis.ramayo@irta.cat

INTRODUCCIÓN

Los animales han evolucionado en estrecha interacción con su microbioma. No obstante, nuestra comprensión sobre los factores que regulan estas interacciones desde una perspectiva hologenómica es aún limitada. Los métodos de secuenciación masiva y la biología computacional han transformado el análisis hologenómico, permitiendo obtener y procesar datos del genoma del hospedador y de sus comunidades microbianas. Sin embargo, estas tecnologías no están exentas de sesgos ya que múltiples factores técnicos pueden influir en la calidad y utilidad de la información generada. El presente estudio tiene como objetivo analizar el impacto de uno de ellos, la profundidad de secuenciación en la caracterización e identificación de biomarcadores en el microbioma asociado al holobionte porcino.

MATERIAL Y MÉTODOS

El estudio se llevó a cabo con 60 cerdos de la raza Duroc asignados durante el periodo de engorde en dos grupos: estrés (n=30) y control (n=30). Se tomaron muestras de heces a los 190 días para generar el perfil de la microbiota (metaT, secuenciación del gen 16S), y el microbioma (metaG, secuenciación del metagenoma completo). En ambos casos, se incluyeron comunidades sintéticas (ZymoBIOMICS®) como controles positivos. A partir de los datos de metaG, se generaron 4 conjuntos que representan diferentes profundidades de secuenciación incluyendo: metaG.sh ($1,0 \times 10^6$), metaG.r1 (25×10^6), metaG.r2 (35×10^6) y metaG.all (60×10^6) lecturas pareadas por grupo. El procesamiento bioinformático se realizó con Qiime2 (Bolyen *et al.*, 2019) para metaT; sylph v0.8.0 (Shaw y Yu, 2024) y nf-core/mag 3.3.0 (Krakau *et al.*, 2022) para generar el perfil taxonómico y ensamblado *de novo* de genomas (MAGs) a partir de metaG. Se comparó la composición taxonómica, similitud entre comunidades, número y calidad de MAGs identificados, evaluando además su capacidad discriminante y para identificar biomarcadores asociados a los grupos experimentales.

RESULTADOS Y DISCUSIÓN

El análisis de comunidades sintéticas no reveló diferencias en la composición y similitud de estas. En cambio, las comunidades porcinas mostraron un importante efecto de la profundidad de secuenciación impactando tanto el número de taxonomías como de MAGs, con rangos que variaron desde 163 géneros, 345 especies y 5 MAGs (metaG.sh) hasta 509 géneros, 1220 especies y 488 MAGs (metaG.all). Por otra parte, la similitud (RV =0,72) entre metaT y metaG sugiere que la diversidad beta derivada de metaT podría servir como un proxy de metaG para estimar la variación interindividual, así como para su inclusión en modelos inferenciales o predictivos. En relación con la capacidad discriminante, observamos que una mayor profundidad de secuenciación tiene un impacto positivo, especialmente a niveles taxonómicos más específicos como el de especies, con una diferencia de precisión del 26,56% (metaG.sh=0,68 vs. metaG.all=0,81).

CONCLUSIÓN

Es necesario aplicar una profundidad de secuenciación adecuada para obtener una representación más completa y precisa de los ecosistemas microbianos, y para identificar biomarcadores. Esto subraya la necesidad de planificar cuidadosamente los experimentos y ajustar este factor técnico en función de los objetivos del estudio y la complejidad del ecosistema microbiano, para optimizar los consumibles, la calidad y utilidad de los datos generados.

REFERENCIAS BIBLIOGRÁFICAS

• Bolyen *et al.* 2019. Nat Bio. • Krakau *et al.* 2022. NAR Geno Bioinfo. • Shaw y Yu. 2024. Nat Bio.

Agradecimientos: Agencia Estatal de Investigación (PID2021-126555OB-I00). YRC es beneficiario de un contrato Ramón y Cajal (RYC2019-027244-I). ACS es beneficiario de un contrato (PRE2022-101829) financiado por el Ministerio de Educación de España. RRL es beneficiaria de un contrato doctoral (2024 FI-1 00034) financiado por AGAUR-FI Joan Oró de la Generalitat de Cataluña.