

Reseña crítica del artículo de investigación “*Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis*”

Presentado por

Carlos Andrés Díaz - código: 202010343

David León - código: 201615216

Cesar Patiño - código: 201924259

En este artículo Jethro Johnson y sus colaboradores presentan en la prestigiosa revista Nature communications en el año 2019 una investigación acerca de la pertinencia del uso del gen completo ARN ribosomal 16S para la clasificación exitosa a nivel de especie y subespecie. Para esto se realizan cuatro experimentos tanto *in silico* como *in vivo* empleando la tecnología de última generación PacBio, esto con el fin de obtener lecturas más largas y contrastarlas con los resultados obtenidos al emplear la tecnología de secuenciación ampliamente usada llamada Illumina que hace uso de regiones hipervariantes del mismo gen para estudios taxonómicos. De esta manera critican algunas de las suposiciones comúnmente realizadas como el uso de las mutaciones de un solo nucleótido (SNP en inglés) para la diferenciación entre especies o el uso del límite de identidad para la agrupación de las secuencias en unidades operacionales taxonómicas (OTU en inglés) en un 97%, entre otros. El artículo resalta el uso de las regiones hipervariantes v1-v3 como regiones con las que se logra buena resolución a nivel de especie, también resalta la incapacidad de la región V4 en dicha labor. Concluye, entre otras cosas, que existen SNPs a nivel intragenómico en la misma especie mostrando el error de la suposición descrita anteriormente, plantea que el uso de OTU está ligado a tecnologías antiguas y de su posible reemplazo por las variantes de secuencias exactas (ESV en inglés) aunque también resalta el error causado por subestimar del primero y de sobreestimar del último.

Si bien el artículo resalta la utilidad de emplear completamente el gen ARN ribosomal 16S sin enfocarse en el uso particular de ninguna plataforma de secuenciación, la justificación que brinda acerca de la contribución a polimorfismos de origen intragenómico que tienen las inserciones y deleciones frente a las sustituciones presentes en las dos subespecies de *E coli* analizadas es poco clara pues no realiza una comparación directa, aprobando el uso de una y descartando el uso de las otras, además ignoran el hecho de que PacBio tiene sensibilidad suficiente para detectar errores a nivel de nucleótidos simples¹. Por otro lado, no discute acerca de las posibles implicaciones que tiene la ampliación del límite de identidad para la agrupación de las secuencias de un 97% al 99% y remarcando superficialmente la relación existente entre el género *bacteroides* y las especies y subespecies descritas por el experimento que hacen significativos los hallazgos reportados.

¹Lang, D., Zhang, S., Ren, P., Liang, F., Sun, Z., & Meng, G. et al. (2020). Comparison of the two up-to-date sequencing technologies for genome assembly: HiFi reads of Pacific Biosciences Sequel II system and ultralong reads of Oxford Nanopore. *Gigascience*, 9(12). doi: 10.1093/gigascience/giaa123