

Evaluación del desempeño de diferentes bases de datos para la clasificación de la familia *Lachnospiraceae*

Sarah R. Delgado-Luna^a, Carolina N. Reséndiz-Nava^a, Gerardo M. Nava-Morales^{a*}

^aFacultad de Química, Universidad Autónoma de Querétaro, Cerro de las Campanas S/N, CP. 76010, Querétaro, México
 *gerardomnava@gmail.com

Palabras clave: *Lachnospiraceae*, clasificación, *16S rRNA*

Introducción

La familia *Lachnospiraceae* pertenece al filo Firmicutes, orden *Clostridiales*; es considerada uno de los principales habitantes del intestino de mamíferos, especialmente en humanos¹. Algunos de los géneros bacterianos asociados a esta familia son considerados marcadores de salud intestinal, como *Blautia* y *Roseburia*²; por su capacidad de producir ácidos grasos de cadena corta (AGCC); también se han relacionado con el desarrollo de enfermedades metabólicas e inflamatorias, como la Obesidad y la Enfermedad de Chron^{3,4}. Debido a la importancia de esta familia bacteriana, la identificación correcta de sus géneros es imprescindible. Por lo anterior, el objetivo del presente trabajo fue evaluar el desempeño de diferentes bases de datos para clasificar correctamente a los diferentes géneros de la familia *Lachnospiraceae*.

Metodología

Se utilizó la base de datos List of Prokaryotic Names with Standing in Nomenclature (LPSN)⁵ para obtener secuencias de referencia del gen *16S rRNA* (~1500 pb) pertenecientes a 57 géneros de la familia *Lachnospiraceae*. Las secuencias fueron comparadas y agrupadas utilizando un valor de corte de 97% de identidad mediante el programa CD-HIT⁶, para reducir el número de secuencias. Las 107 secuencias representativas obtenidas se sometieron a un proceso de clasificación utilizando cuatro bases de datos a través del programa SINA de SILVA⁷ con un mínimo de identidad de 0.97 y un rechazo de secuencias con identidad menor al 70%. Las cuatro bases de datos seleccionadas para evaluar su desempeño en la clasificación fueron: Genome Taxonomy Database (GTDB)⁸, "The All-Species Living Tree" Project (LTP)⁹, Ribosomal Database Project (RDP)¹⁰ y SILVA¹¹.

Resultados y discusión

Se observaron importantes diferencias en la clasificación de secuencias de la familia *Lachnospiraceae*. Las cuatro bases de datos utilizadas difirieron en la clasificación de secuencias a nivel de orden, donde una alta proporción (100%) de secuencias fueron "no clasificadas"; asimismo, únicamente tres géneros coincidieron a nivel de familia. GTDB reportó el mayor número de secuencias no clasificadas con un 84.74%, siguiendo RDP (62.71%), LTP (50.84%) y SILVA (45.76%). La base de datos LTP clasificó correctamente el 82.75% de las secuencias de referencia, seguido de SILVA (71.87%), GTDB (77.7%) y RDP (65.21%).

La base de datos con menor cantidad de conflictos taxonómicos a nivel de género fue LTP con un 17.24%. Asimismo, los errores observados en la clasificación denotan la falta de actualización en estas bases, ya que se identificaron géneros que han sido reclasificados, tal es el caso de *Eubacterium*¹², que su especie -

rectalis- fue catalogada como *Agathobacter rectalis*; además, secuencias anotadas como *Clostridium*¹³, fueron reasignadas a otros géneros como *Enterocloster*, *Lacrimispora* o *Anaerocolumna*. SILVA clasificó géneros no validados como *Lachnoclostridium*¹⁴ en secuencias de referencia pertenecientes a géneros *Enterocloster*, *Mediterraneibacter* y *Merdimonas*.

Conclusiones

Para el análisis de la familia *Lachnospiraceae*, las mejores bases de datos para su clasificación son LTP y SILVA. La constante reclasificación de géneros obliga a los investigadores que trabajan en estudios del microbioma, a actualizarse constantemente siguiendo los marcos de referencia determinados por el International Journal of Systematic and Evolutionary Microbiology (IJSEM) y verificando el nivel de actualización de las diferentes bases de datos.

Referencias

- (1) Meehan, C. J.; Beiko, R. G. *Genome Biology and Evolution* **2014**, 6 (3), 703–713.
- (2) Vacca, M.; Celano, G.; Calabrese, F. M.; Portincasa, P.; Gobetti, M.; De Angelis, M. *Microorganisms* **2020**, 8 (4), 573.
- (3) Humbel, F.; Rieder, J. H.; Franc, Y.; Juillerat, P.; Scharl, M.; Misselwitz, B.; Schreiner, P.; Begeré, S.; Rogler, G.; von Känel, R.; Yilmaz, B.; Biedermann, L. *Clinical Gastroenterology and Hepatology* **2020**, 18 (9), 2019–2029.e11.
- (4) Kameyama, K.; Itoh, K. *Microb. Environ.* **2014**, 29(4), 427–430.
- (5) Parte, A. C. *International Journal of Systematic and Evolutionary Microbiology*, **2018**, 68 (6), 1825–1829.
- (6) Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. *Bioinformatics* **2012**, 28 (23), 3150–3152.
- (7) Pruesse, E.; Peplies, J.; Glöckner, F. O. *Bioinformatics* **2012**, 28 (14), 1823–1829.
- (8) Chaumeil, P.-A.; Mussig, A. J.; Hugenholtz, P.; Parks, D. H. *Bioinformatics* **2020**, 36 (6), 1925–1927.
- (9) Yarza, P.; Richter, M.; Peplies, J.; Euzéby, J.; Amann, R.; Schleifer, K.-H.; Ludwig, W.; Glöckner, F. O.; Rosselló-Móra, R. *Syst. Appl. Microbiol.* **2008**, 31 (4), 241–250.
- (10) Wang, Q.; Garrity, G. M.; Tiedje, J. M.; Cole, J. R. *Appl. Environ. Microbiol.* **2007**, 73 (16), 5261–5267.
- (11) Yilmaz, P.; Parfrey, L. W.; Yarza, P.; Gerken, J.; Pruesse, E.; Quast, C.; Schweer, T.; Peplies, J.; Ludwig, W.; Glöckner, F. O. *Nucleic Acids Res* **2014**, 42 (D1), D643–D648.
- (12) Rosero, J. A.; Killer, J.; Sechovcová, H.; Mrázek, J.; Benada, O.; Fliegerová, K.; Havlík, J.; Kopečný, J. *Int. J. Syst. Evol. Microbiol.* **2016**, 66 (2), 768–773.
- (13) Haas, K. N.; Blanchard, J. L. *Int. J. Syst. Evol. Microbiol.* **2020**, 70 (1), 23–34.
- (14) Brahimi, S.; Cadoret, F.; Fournier, P.-E.; Moal, V.; Raoult, D. *New Microbes New Infect* **2017**, 16, 73–75.