Towards global multi-host bacterial genomic epidemiology

Hacia una epidemiología genómica bacteriana global y multihospedador

Santiago Castillo-Ramírez[1, a]

[1] Programa de Genómica Evolutiva, Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, México

SOUTH HEALTH

Citar como / Cite as: Castillo-Ramírez S. Towards global multi-host bacterial genomic epidemiology. South Health. 2026;3:e004. doi: 10.21142/SH-03-2026-e004

Fecha de recibido / Reception: 25/05/2026
Fecha de aceptado / Accept: 04/08/2026
Revisado por pares / Peer reviewed

Editor: Joaquim Ruiz

*Corresponding Author:
Santiago Castillo-Ramírez
iago@ccg.unam.mx

Abstract

Whole Genome Sequencing (WGS) is now the gold standard for high-resolution molecular epidemiology of bacteria. There are several approaches to analyse WGS data, from genome phylogenies or core genome MLST schemes, to Average Nucleotide Identity analysis, Pangenome analysis, and even to molecular dating analysis. These approaches have advanced the genomic epidemiology of bacterial pathogens thoroughly. Yet, a few important sampling challenges remain. First, the strong bias towards clinical surveillance, yet many of the so-called "human pathogens" are also found in other hosts. Secondly, the very patchy distribution of countries with the proper funding to conduct WGS, with many countries having hardly any isolates sequenced. Finally, many studies just focus on antibiotic-resistant isolates, overlooking the diversity hosted by the antibiotic-susceptible isolates. Thus, ideally, future studies should conduct global multi-host bacterial genomic epidemiology despite the antibiotic phenotype of the isolates.

Keywords: Genomic epidemiology, ESKAPE pathogens, genotyping, molecular epidemiology, One Health, LMICs

Resumen

La secuenciación del genoma completo (WGS, por sus siglas en inglés) constituye actualmente el estándar de oro para la epidemiología molecular bacteriana de alta resolución. Existen diversos enfoques para analizar los datos de WGS, que van desde las filogenias genómicas o los esquemas de tipificación de secuencias multilocus (MLST) del genoma central (core genome en inglés), hasta el análisis de la Identidad Promedio de Nucleótidos (ANI, por sus siglas en ingles), el análisis del pangenoma e incluso el análisis de datación molecular. Estos métodos han impulsado considerablemente la epidemiología genómica de los patógenos bacterianos. No obstante, persisten algunos desafíos importantes en materia de muestreo. En primer lugar, existe un marcado sesgo hacia la vigilancia clínica, a pesar de que muchos de los denominados "patógenos humanos" se encuentran también en otros hospedadores. En segundo lugar, la distribución es muy desigual en cuanto a los países que cuentan con la financiación adecuada para realizar WGS, ya que en muchos de ellos apenas se han secuenciado aislados. Por último, muchos estudios se centran exclusivamente en aislados resistentes a los antibióticos, pasando por alto la diversidad presente en los aislados sensibles a los mismos. Por consiguiente, lo ideal sería que los estudios futuros abordaran la epidemiología genómica bacteriana a escala global y abarcando múltiples hospedadores, independientemente del fenotipo de resistencia a los antibióticos de los aislados.

Palabras Clave: Epidemiología genómica, Patógenos del grupo ESKAPE, Genotipado Molécular genotyping, Epidemiología molecular, Una Salud, LMICs.

Main text

Whole Genome Sequencing (WGS) has significantly moved forward the field of bacterial molecular epidemiology. Clearly, the high level of genetic resolution of WGS has made it the most suitable option for conducting bacterial molecular epidemiology. Currently, there are several approaches to make sense of the WGS data and determine the genomic epidemiology of the pathogen under study. The two most used approaches for analysing WGS data of (pathogenic) bacteria are core phylogenies and core genome MLST (cgMLST) schemes. The former consists of constructing phylogenies on the core genome or some part of it. Where the core genome is all those genomic regions present in all the isolates/taxa/OTUs considered. This strategy has been very useful for inferring transmission dynamics at very different scales, both at temporal and geographic dimensions. For instance, recently, the transmission dynamics of a polyclonal dissemination of *Acinetobacter baumannii* in a Paraguayan hospital were established (Bello-Lopez *et al*., 2025). Considering a broader geographic scale, in another study, a core phylogenomic analysis was employed to determine the dissemination of a high-risk clone of *Pseudomonas aeruginosa* circulating in Chilean hospitals (Opazo-Capurro *et al*., 2024). The second most employed approach, cgMLST schemes, is an extension of the traditional MLST, which scales to the genome level, and instead of considering just 7 loci, the idea is to consider the loci belonging to the core genome. Similar to the former approach, cgMLST has been very productive in studying the dissemination of bacterial pathogens at different scales. This strategy was very useful to study the genetic relationships of extensively drug-resistant isolates of *A. baumannii* in a Mexican hospital (Moreno-Manjón *et al*., 2023). At the other extreme, i. e., going beyond the species level, a recent cgMLST scheme has been implemented to determine lineages of the species complex *Borrelia burgdoferi* sensu lato (Hepner *et al*., 2025). In addition, some researchers also use population structure analyses to delineate bacterial populations. For example, a population structure analysis was essential to infer the trans-Atlantic transmission of the Lyme disease agent *Borrelia burgdorferi* sensu stricto (Castillo-Ramírez *et al*. 2025). Very recently, Average Nucleotide Identity (ANI) based definitions of intra-species units have been implemented for bacteria (Castillo-Ramírez, 2024; Raghuram *et al*., 2024; Viver *et al*., 2024). Some studies have even conducted molecular dating analysis to establish the emergence of clones or lineages for important pathogens, like *A. baumannii* and *Staphylococcus aureus* (Challagundla *et al*., 2018; Frisch *et al*., 2018; Graña-Miraglia *et al*., 2020; Mateo-Estrada *et al*., 2021). For instance, a few years ago, a study dated the introduction of several important lineages of *A. baumannii* in a Mexican hospital (Mateo-Estrada *et al*., 2021). Finally, pangenome analyses have been extremely useful to fully appreciate the diversity of gene families (Tettelin and Medini, 2020), not least those related with virulence and/or resistance phenotypes. In this regard, a recent study paying attention to the pangenome of non-human isolates of *A. baumannii* shows that these isolates have vast diversity of antibiotic resistance genes (Aguilar-Vera *et al*., 2026).

Clearly, bacterial genomic epidemiology is coming of age. Yet, some important issues must still be addressed. A major issue is the slanted sampling, where most of our surveillance is in (human) clinical settings. Most genomic epidemiology studies only consider human (clinical) isolates. It is important to mention that many bacteria can dwell in several hosts (or be found in different sources). Actually, several ESKAPE pathogens are well known for their ability to inhabit multiple hosts (Castillo-Ramírez *et al*., 2025; Muloi *et al*., 2022; Roy *et al*., 2024). Hence, a multi-host genomic epidemiology might be a more natural way to try to understand the transmission of many of the reputed "human pathogens". Of note, this completely goes along with the One Health ethos, which considers that the health of humans is interconnected with both animal and environmental health. Given the underdeveloped environmental and human health infrastructure in some low- and middle-income countries (LMICs), they are natural places to really carry out One Health surveillance, and this type of surveillance should be enhanced. However, rather than delegating the leading roles of the future multi-host epidemiology studies to principal investigators (PIs) from the Global North, we must ensure that the PIs from LMICs lead such projects and avoid parachute science. This would undoubtedly strengthen science in those countries, curtailing dependence on scientific institutions in the Global North. Another issue affecting many genomic epidemiology studies for many bacteria is that, due to a lack of funding and resources, there is a patchy sampling distribution, considering the geographic origin of the isolates included. Clearly, LMICs also need to play a vital role in extending the geographic sampling of many pathogens. Here again, we must ensure that the PIs from LMICs have the lead roles to avoid colonial science. Finally, many genomic epidemiology studies concentrate on just antibiotic-resistant isolates, which is an oversight, as this overlooks the diversity within the antibiotic-susceptible isolates. In this context, it is relevant to mention that antibiotic-susceptible isolates can be important carriers of not only mobile genetic elements but also virulence factors, which also contribute to the evolution and persistence of bacterial pathogens. Thus, this type of isolates should also be included in our genomic surveillance efforts. There is no doubt that LMICs are bound to play a significant role in approaching the global multi-host genomic epidemiology of the ESKAPE pathogens.

In conclusion, getting worldwide and multi-host WGS data sets is bound to improve our inferences of the transmission dynamics of many bacteria, pathogenic and non-pathogenic alike. However, we really want to make sure that LMICs play a major role in this endeavour.

Funding

SCR is supported by a PAPIIT grant (number IN209524) from UNAM.

Author Contribution

SCR conceived and wrote this comment.

Conflicts of interest

The author declares that there are no conflicts of interest.

Transparency declaration

None to declare.

REFERENCES

Aguilar-Vera A, López-Sánchez R, Hernández-Alvarez AJ, Vargas-Peralta D, Peralta H, Castillo-Ramírez S. Global multi-host genomic epidemiology of *Acinetobacter baumannii* reveals transmission at one health interfaces. Nat Commun. 2026;17:5832. doi: 10.1038/s41467-026-72585-4

Bello-López E, Kawabata A, Cantero J, Mendoza S, Pertile E, Perez-Osegura A, *et al*. Genomic epidemiology reveals antibiotic resistance transfer and polyclonal dissemination of *Acinetobacter baumannii* in a Paraguayan hospital. Antimicrob Agents Chemother. 2025;69:e0007725. doi: 10.1128/aac.00077-25

Castillo-Ramírez S. On the road to genomically defining bacterial intra-species units. mSystems. 2024;9:e0058424. doi: 10.1128/msystems.00584-24

Castillo-Ramírez S, Aguilar-Vera A, Kumar A, Evans B. *Acinetobacter baumannii*: much more than a human pathogen. Antimicrob Agents Chemother. 2025;69:e0080125. doi: 10.1128/aac.00801-25

Castillo-Ramírez S, Fingerle V, Jungnick S, Straubinger RK, Krebs S, Blum H, *et al*. Trans-Atlantic exchanges have shaped the population structure of the Lyme disease agent *Borrelia burgdorferi* sensu stricto. Sci Rep. 2016;6:22794. doi: 10.1038/srep22794

Challagundla L, Reyes J, Rafiqullah I, Sordelli DO, Echaniz-Aviles G, Velazquez-Meza ME, *et al*. Phylogenomic classification and the evolution of clonal complex 5 methicillin-Resistant *Staphylococcus aureus* in the western hemisphere. Front Microbiol. 2018;9:1901. doi: 10.3389/fmicb.2018.01901

Frisch MB, Castillo-Ramírez S, Petit RA, III, Farley MM, Ray SM, Albrecht VS, *et al*. Invasive methicillin-resistant *Staphylococcus aureus* USA500 strains from the U.S. Emerging Infections Program constitute three geographically distinct lineages. mSphere. 2018;3:e00571-17. doi: https://doi.org/10.1128/mSphere.00571-17.

Graña-Miraglia L, Evans BA, López-Jácome LE, Hernández-Durán M, Colín-Castro CA, *et al*. Origin of OXA-23 variant OXA-239 from a recently emerged lineage of *Acinetobacter baumannii* international clone V. mSphere. 2020;5:e00801-19. doi: https://doi.org/10.1128/mSphere.00801-19

Hepner S, Jolley KA, Castillo-Ramirez S, Mourkas E, Dangel A, Wieser A, *et al*. A core genome MLST scheme for *Borrelia burgdorferi* sensu lato improves insights into the evolutionary history of the species complex. Cell Rep Methods. 2025;5:100935. doi: 10.1016/j.crmeth.2024.100935

Mateo-Estrada V, Fernández-Vázquez JL, Moreno-Manjón J, Hernández-González IL, Rodríguez-Noriega E, Morfín-Otero R, *et al*. Accessory genomic epidemiology of cocirculating *Acinetobacter baumannii* clones. mSystems. 2021;6:e0062621. doi: 10.1128/mSystems.00626-21

Moreno-Manjón J, Castillo-Ramírez S, Jolley KA, Maiden MCJ, Gayosso-Vázquez C, Fernández-Vázquez JL, *et al*. *Acinetobacter baumannii* IC2 and IC5 isolates with co-existing $bla_{OXA-143-like}$ and bla_{OXA-72} and exhibiting strong biofilm formation in a Mexican hospital. Microorganisms. 2023;11:2316. doi: 10.3390/microorganisms11092316

Muloi DM, Hassell JM, Wee BA, Ward MJ, Bettridge JM, Kivali V, *et al*. Genomic epidemiology of *Escherichia coli*: antimicrobial resistance through a One Health lens in sympatric humans, livestock and peri-domestic wildlife in Nairobi, Kenya. BMC Med. 2022;20:471. doi: 10.1186/s12916-022-02677-7

Opazo-Capurro A, Aguilar-Vera OA, González-Muñoz P, Amsteins-Romero L, Quiroga M, Encina A, *et al*. Genomic and phylogenomic characterization of carbapenem-resistant *Pseudomonas aeruginosa* 'high-risk' clone O4/ExoS+/ST654 circulating in Chilean hospitals. J Glob Antimicrob Resist. 2024;38:205-211. doi: 10.1016/j.jgar.2024.05.015

Raghuram V, Petit RA, Karol Z, Mehta R, Weissman DB, Read TD. Average nucleotide identity-based *Staphylococcus aureus* strain grouping allows identification of strain-specific genes in the pangenome. mSystems 2024;9:e00143-24. doi: 10.1128/msystems.00143-24

Roy MC, Chowdhury T, Hossain MT, Hasan MM, Zahran E, Rahman MM, et al. Zoonotic linkage and environmental contamination of methicillin-resistant *Staphylococcus aureus* (MRSA) in dairy farms: A one health perspective. One Health. 2024;18:100680. doi: 10.1016/j.onehlt.2024.100680

Tettelin H, Medini D (eds). The pangenome: diversity, dynamics and evolution of genomes. Springer, Cham, Switzerland. 2020. doi: 10.1007/978-3-030-38281-0

Viver T, Conrad RE, Rodriguez-R LM, Ramírez AS, Venter SN, Rocha-Cárdenas J, *et al*. Towards estimating the number of strains that make up a natural bacterial population. Nat Commun. 2024;15:544. doi: 10.1038/s41467-023-44622-z