

Supplementary File 1

Viral sequencing to inform the global elimination of dog-mediated rabies - a systematic review

Gurdeep Jaswant^{1,2,3,4,#}, Criselda T. Bautista^{2,5,#}, Brian Ogoti^{1,6}, Joel Changalucha^{4,7}, Julius O. Oyugi^{1,8}, Kathryn Campbell², Mumbua Mutunga^{1,6}, S.M. Thumbi^{1,6,7,8,9,10,#}, Katie Hampson^{2,#}, Kirstyn Brunker^{2,#}

¹Faculty of Health Sciences, Institute of Tropical and Infectious Diseases, University of Nairobi, Nairobi 00202, Kenya.

²School of Biodiversity, One Health & Veterinary Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, UK.

³Food Biotechnology and Microbiology, Tanzania Industrial Research Development Organization, Dar es Salaam 14111, Tanzania.

⁴Environmental Health and Ecological Sciences, Ifakara Health Institute, Dar es Salaam Plot 463, Tanzania.

⁵Research Institute for Tropical Medicine, Muntinlupa City, Metro Manila 1781, Philippines.

⁶Faculty of Health Sciences, Center for Epidemiological Modelling and Analysis, University of Nairobi, Nairobi 00202, Kenya.

⁷College of Veterinary Medicine & Biomedical Science, Sokoine University of Agriculture, Morogoro 67804, Tanzania.

⁸Department of Medical Microbiology, Faculty of Health Sciences, University of Nairobi, Nairobi 00202, Kenya.

⁹Institute of Immunology & Infection Research, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FL, Scotland, UK.

¹⁰Paul G Allen School for Global Health, Washington State University, Pullman, WA 99164, USA.

[#]Authors contributed equally.

Correspondence to: Dr. Kirstyn Brunker, Graham Kerr building, School of Biodiversity, One Health & Veterinary Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, UK. E-mail: kirstyn.brunker@glasgow.ac.uk

Abstract

Background: Rabies is a deadly yet neglected infectious disease. Present in almost 150 countries around the world, with most deaths reported in Asia and Africa, rabies is a serious pressing issue worldwide. A global strategy has been initiated with the aim of eliminating human deaths from rabies spread by domestic dogs by 2030 ('Zero by 30'). Genomic surveillance is a tool that can potentially support the 'Zero by 30' strategy.

Methods: The databases PubMed, Google Scholar and Web Of Science will be searched to identify original studies published since the year 2000 with the following search terms 'Rabies AND (genom* OR sequenc* OR phylo* OR molecular) AND (control OR surveillance OR eliminat*)'. Pre-defined inclusion and exclusion criteria will be used to select relevant studies and the selection procedure will be shown by a Preferred Reporting Items for Systematic reviews and Meta-analysis study flow diagram (PRISMA). Data will be extracted including author, year of publication, location of study, study design, sequencing platform, and coverage of genome (whole or partial), type and number of samples sequenced, infected host species, analysis methods, conclusion(s) of the study and any recommendations for control measures or surveillance derived from the genomic data. Data will be summarized in terms of trends in published papers, geographical coverage, sequencing platforms, length, and available genomic data.

Expected output: To get geographical coverage of the country with sequencing data in Africa to see where gaps are to expand the need of genomic surveillance. To give recommendations on how to improve genomic surveillance based on sample types, sequencing platforms, and data management. Key messages from study's conclusion regarding how genomic surveillance provides insights to inform Zero by 30 by drawing the unique message from genomic data.

Registration: This protocol will be submitted to the PROSPERO database for registration.

Keywords: Sequencing, phylogenetic, lyssavirus, molecular techniques

Background

Rabies Virus (RABV) poses the greatest public health threat, causing an estimated 60,000 deaths annually, almost all of which occur in Low- and Middle-Income Countries (LMICs) [1]. RABV is most commonly transmitted through bites from infected hosts in the orders *Chiroptera* and *Carnivora* [2]. Domestic dogs are the main source of transmission to humans, but as a multi-host pathogen, wild carnivores also serve as primary RABV hosts with host-associated variants recorded in certain geographies [3]. For example, wildlife such as raccoons, skunk, and foxes each maintain different RABV variants in localities across North America [3]. Generally, RABV is referred to according to these host-associated variants (sometimes termed biotype, see definitions introduced in Box 1). Phylogenetic analysis enables further classification of RABV diversity into clades, subclades, and lineages, usually associated with specific geographic areas and/or hosts. The RABV genome is 12 kilobases (kb) in length [4], comprising five genes encoding the nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and the large polymerase protein (L) [5]. Like other RNA viruses, RABV exhibits elevated mutation rates because of the absence of proofreading activity in the L protein [6]. Viral sequence data is informative because of these elevated mutation rates, which generate genetic diversity enabling improved tracking of viral spread and understanding of viral dynamics across space and over time.

There is no treatment for rabies once clinical signs begin, but post-exposure prophylaxis (PEP) administered shortly after a rabies exposure is almost 100% effective in preventing the fatal onset of disease [7]. Canine rabies elimination is possible through mass dog vaccination, as demonstrated in Europe, North America, parts of Asia and much of Latin America [8]. Several countries where dog-mediated rabies was previously endemic have now been declared rabies-free (Western Europe, Canada, the USA, and Japan) or are approaching elimination as a result of sustained dog vaccination [8]. According to the World Health Organisation (WHO), to eliminate dog-mediated rabies, vaccination campaigns need to achieve coverage of at least 70% of the dog population and be conducted annually for at least three years [9]. The incidence of rabies in Latin America has declined dramatically over recent decades due to coordinated regional elimination programs underpinned by this approach [8]. In contrast, most LMICs in Asia and Africa have not allocated sufficient budget to control this neglected disease. In these endemic countries, rabies

surveillance is typically poor and challenges to rabies control include lack of understanding of dog ownership patterns, dog population sizes and dog accessibility for vaccination as well as cultural practices including dog meat consumption [10]. To address these challenges, international organisations recently joined forces under the United Against Rabies collaboration to advocate for the global goal of ‘Zero by 30’, to end human deaths from dog-mediated rabies by 2030 [11].

Surveillance plays a critical role in the control and elimination of infectious diseases [12]. Surveillance entails the continuous, systematic collection, analysis, interpretation, and timely dissemination of health-related information [13], serving as the foundation for planning, execution and evaluation of public health strategies. For instance, surveillance aids in producing data on the effectiveness of interventions, thus offering valuable insights for decision-making crucial for elimination initiatives like 'Zero by 30' [14]. Increasingly, surveillance involves genetic data, for pathogen diagnosis, for determining risks associated with a pathogen or its susceptibility to drugs, as well as to identify the source of outbreaks and to characterise pathogen spread [12]. Linked with locations, pathogen genetic data have uncovered different aspects of disease movement, from global migration dynamics to local transmission pathways for pathogens such as Influenza [15,16], Ebola [17], Zika [18], Yellow fever [19,20], Mpox [21–24] and SARS-CoV-2 [25]. Sequencing approaches have potential to enhance rabies surveillance and provide actionable information to inform rabies control programs locally and as part of ‘Zero by 30’. For example, viral sequence data can distinguish continuous undetected local circulation from incursions and potentially identify their sources [26]. More generally, sequencing could provide key insights into how rabies circulates within different populations and the processes responsible for RABV maintenance in specific localities [12,27].

Use of pathogen sequence data within surveillance programmes is, however, not yet routine in most LMICs. Constraints include lack of local sequencing capacity, trained personnel and laboratory resources, affected by the costs of and access to reagents and consumables, as well as power supplies and cold chain [17]. Sequencing technologies have become more affordable, and efforts are underway to improve their accessibility [28]. Indeed, growth in sequencing capacity in LMICs during the COVID-19 pandemic provided evidence of the feasibility of scaling up

molecular diagnostics, but also highlighted operational challenges. For example, public health laboratories in Nigeria capable of molecular identification of SARS-COV-19 from clinical specimens increased from four to 72 laboratories in 2020 [29]. In this systematic review, our goal will be to examine the current extent of the application of genetic approaches to RABV surveillance globally and how, going forward, these approaches can contribute to the global strategy to eliminate human deaths from dog-mediated rabies.

Methods

Search strategy

A systematic search will be done on PubMed, Web of Science and Google Scholar electronic databases to identify original studies that reported genomic surveillance of rabies to support rabies elimination. Advanced searches with Boolean operators and quotations will be performed using the following key terms: ‘rabies AND (genom* OR sequenc* OR phylo* OR molecular) AND (control OR surveillance OR eliminat*)’. Further manual searches will be performed for additional relevant studies.

Selection of studies

Data will be extracted from any design (prospective/ retrospective). Pre-defined inclusion and exclusion criteria will be used to select relevant studies and the selection procedure will be shown by a Preferred Reporting Items for Systematic reviews and Meta-analysis study flow diagram (PRISMA).

Inclusion and exclusion criteria

To ensure that relevant studies are included the following inclusion criteria will be used for screening: studies must address either canine rabies, human rabies or terrestrial wildlife rabies (i.e. not bat rabies), and use molecular techniques with sequencing data either for diagnosis or surveillance of rabies. We will exclude studies reported as literature reviews without presenting data, studies that are not published in English language, duplicated papers, that do not focus on rabies or include genomic/ sequencing data. We will follow PRISMA (Moher et al. 2009) guidelines to determine the Population, Intervention, Comparison and Outcome of the study (PICO), which for our study covers:

P (Population) = Rabies virus

I (Intervention) = genomic sequencing approaches

C (Comparison) = Known rabies control and prevention measures such as Mass dog vaccination and Post Exposure Prophylaxis (PEP)

O (Outcome) = Primary outcome - rabies control guidance; Secondary outcome – Other message from genomic surveillance

Management of identified articles and Quality assessment

All the articles identified from database searches will be exported for duplicate removal, screening of titles, abstracts and eligibility assessment according to the specified inclusion and exclusion criteria. Two independent reviewers will assess the quality of studies to be included in the systematic review. Any discrepancy observed between reviewers regarding the quality of selected study (s) will be resolved through discussion.

Data extraction and analysis

Data from eligible studies will be extracted into spreadsheets. with the following information , author and year of publication; location of study (country and subnational administrative unit if reported), study design, platform for sequencing, type of samples used (brain, saliva, vaccine or other), species of infected animal host (domestic dog, wildlife or other domestic animal, indicating the species involved), sample size (n), data analysis methods, conclusion(s) of the study and any control measure derived from the sequencing data.

The extracted data will be used to summarise the number of published papers and trends of time, the geographical coverage of the article the (richness of the genomic data available from different areas) , what were the most commonly used sequencing platforms, and implications of the studies on how genomic data draws insight for rabies control and elimination.

Expected output

To get geographical coverage of the country with sequencing data in Africa so as to see where gaps are to expand the need of genomic surveillance. To give recommendations on how to improve genomic surveillance based on sample types, sequencing platforms, and data

management. Key messages from studies conclusion regarding how genomic surveillance provides insights to inform Zero by 30 by drawing the unique message from genomic data.

Ethical approval and consent to participate

Not applicable

Consent for publication

The authors consented for publication

Funding

Wellcome funding through a Senior Research Fellowship to KH [207569], a Training Fellowship in Public Health and Tropical Medicine to SMT [110330] and Institutional Strategic Support Fund grants to KB at the University of Glasgow [204820]. National Institute for Health Research - Genomics and Modelling to the Control of Virus Pathogens (GeMVi) fellowship funded by National Institute for Health Research (NIHR) to GJ [176382]

REFERENCE

1. Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, et al. Estimating the Global Burden of Endemic Canine Rabies. *PLoS Neglected Tropical Diseases*. 2015;9: 1–20. doi:10.1371/journal.pntd.0003709
2. Worsley-Tonks KEL, Escobar LE, Biek R, Castaneda-Guzman M, Craft ME, Streicker DG, et al. Using host traits to predict reservoir host species of rabies virus. *PLoS Negl Trop Dis*. 2020;14: e0008940. doi:10.1371/journal.pntd.0008940
3. WHO. Epidemiology of Rabies | Rabies - Bulletin - Europe. [cited 31 Jul 2023]. Available: <https://www.who-rabies-bulletin.org/site-page/epidemiology-rabies>
4. Pant GR, Lavenir R, Wong FYK, Certoma A, Larrous F, Bhatta DR, et al. Recent Emergence and Spread of an Arctic-Related Phylogenetic Lineage of Rabies Virus in Nepal. Munoz-Zanzi C, editor. *PLoS Negl Trop Dis*. 2013;7: e2560. doi:10.1371/journal.pntd.0002560
5. Wunner WH, Larson JK, Dietzschold B, Smith CL. The Molecular Biology of Rabies Viruses. *Reviews of Infectious Diseases*. 1988;10: S771–S784. doi:10.1093/clinids/10.Supplement_4.S771
6. Voloch CM, Capellão RT, Mello B, Schrago CG. Analysis of Adaptive Evolution in Lyssavirus Genomes Reveals Pervasive Diversifying Selection during Species Diversification. *Viruses*. 2014;6: 4465–4478. doi:10.3390/v6114465
7. States M, Strategic WHO, Group A, Grade T, Sage T. Weekly epidemiological record Relevé épidémiologique hebdomadaire. 2018; 201–220.
8. WHO. WHO Expert Consultation on Rabies. In: WHO. (2005) ‘WHO Expert Consultation on Rabies’, World Health Organization technical report series, 931: 1–121. [Internet]. 2005 [cited 10 Jun 2021]. Available: https://apps.who.int/iris/bitstream/handle/10665/85346/9789241209823_eng.pdf?sequence=1
9. WHO Expert Consultation on Rabies.
10. Nguyen AKT, Vu AH, Nguyen TT, Nguyen DV, Ngo GC, Pham TQ, et al. Risk Factors and Protective Immunity Against Rabies in Unvaccinated Butchers Working at Dog Slaughterhouses in Northern Vietnam. *Am J Trop Med Hyg*. 2021;105: 788–793. doi:10.4269/ajtmh.20-1172
11. Minghui R, Stone M, Semedo MH, Nel L. New global strategic plan to eliminate dog-mediated rabies by 2030. *The Lancet Global Health*. 2018;6: e828–e829. doi:10.1016/S2214-

109X(18)30302-4

12. Talbi C, Lemey P, Suchard MA, Abdelatif E, Elharrak M, Jalal N, et al. Phylodynamics and Human-Mediated Dispersal of a Zoonotic Virus. *PLOS Pathogens*. 2010;6: e1001166. doi:10.1371/journal.ppat.1001166
13. Cutts FT, Waldman RJ, Zoffman HMD. Surveillance for the Expanded Programme on Immunization. 1993;71: 7.
14. THACKER SB, BERKELMAN RL. PUBLIC HEALTH SURVEILLANCE IN THE UNITED STATES. *Epidemiologic Reviews*. 1988;10: 164–190. doi:10.1093/oxfordjournals.epirev.a036021
15. Bedford T, Cobey S, Beerli P, Pascual M. Global Migration Dynamics Underlie Evolution and Persistence of Human Influenza A (H3N2). *PLOS Pathogens*. 2010;6: e1000918. doi:10.1371/journal.ppat.1000918
16. Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, et al. Unifying Viral Genetics and Human Transportation Data to Predict the Global Transmission Dynamics of Human Influenza H3N2. *PLOS Pathogens*. 2014;10: e1003932. doi:10.1371/journal.ppat.1003932
17. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature*. 2016;530: 228–232. doi:10.1038/nature16996
18. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols*. 2017;12: 1261–1266. doi:10.1038/nprot.2017.066
19. Giovanetti M, Pinotti F, Zanluca C, Fonseca V, Nakase T, Koishi AC, et al. Genomic epidemiology unveils the dynamics and spatial corridor behind the Yellow Fever virus outbreak in Southern Brazil. *Science Advances*. 2023;9: eadg9204. doi:10.1126/sciadv.adg9204
20. Faria NR, Kraemer MUG, Hill SC, Goes de Jesus J, de Aguiar RS, Iani FCM, et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science*. 2018;361: 894–899. doi:10.1126/science.aat7115
21. Roychoudhury P, Sereewit J, Xie H, Nunley E, Bakhsh SM, Lieberman NAP, et al. Genomic Analysis of Early Monkeypox Virus Outbreak Strains, Washington, USA. *Emerg Infect Dis*. 2023;29: 644–646. doi:10.3201/eid2903.221446

22. Kugelman JR, Johnston SC, Mulembakani PM, Kisalu N, Lee MS, Koroleva G, et al. Genomic Variability of Monkeypox Virus among Humans, Democratic Republic of the Congo. *Emerg Infect Dis.* 2014;20: 232–239. doi:10.3201/eid2002.130118
23. Forni D, Cagliani R, Molteni C, Clerici M, Sironi M. Monkeypox virus: The changing facets of a zoonotic pathogen. *Infection, Genetics and Evolution.* 2022;105: 105372. doi:10.1016/j.meegid.2022.105372
24. Li Y, Hou J, Sun Z, Hu J, Thilakavathy K, Wang Y, et al. Monkeypox virus 2022, gene heterogeneity and protein polymorphism. *Sig Transduct Target Ther.* 2023;8: 1–4. doi:10.1038/s41392-023-01540-2
25. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido D da S, Mishra S, et al. Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv.* 2021; 2021.02.26.21252554. doi:10.1101/2021.02.26.21252554
26. Trewby H, Nadin-Davis SA, Real LA, Biek R. Processes Underlying Rabies Virus Incursions across US–Canada Border as Revealed by Whole-Genome Phylogeography. *Emerg Infect Dis.* 2017;23: 1454–1461. doi:10.3201/eid2309.170325
27. Layan M, Dellicour S, Baele G, Cauchemez S, Bourhy H. Mathematical modelling and phylodynamics for the study of dog rabies dynamics and control: A scoping review. *PLOS Neglected Tropical Diseases.* 2021;15: e0009449. doi:10.1371/journal.pntd.0009449
28. Brunker K, Jaswant G, Thumbi SM, Lushasi K, Lugelo A, Czupryna AM, et al. Rapid in-country sequencing of whole virus genomes to inform rabies elimination programmes. *Wellcome Open Res.* 2020;5: 3. doi:10.12688/wellcomeopenres.15518.2
29. Okeke IN, Ihekweazu C. The importance of molecular diagnostics for infectious diseases in low-resource settings. *Nat Rev Microbiol.* 2021;19: 547–548. doi:10.1038/s41579-021-00598-