Archival Collections are Important in the Study of the Biology, Diversity, and Evolution of Arboviruses

Alyssa T. Pyke and David Warrilow

Public Health Virology Laboratory, Forensic and Scientific Services, Queensland Health, Archerfield, Queensland, Australia.

Supplementary Issue: Evolution of the Insect Virome

ABSTRACT: Historically, classifications of arboviruses were based on serological techniques. Hence, collections of arbovirus isolates have been central to this process by providing the antigenic reagents for these methods. However, with increasing concern about biosafety and security, the introduction of molecular biology techniques has led to greater emphasis on the storage of nucleic acid sequence data over the maintenance of archival material. In this commentary, we provide examples of where archival collections provide an important source of genetic material to assist in confirming the authenticity of reference strains and vaccine stocks, to clarify taxonomic relationships particularly when isolates of the same virus species have been collected across a wide expanse of time and space, for future phenotypic analysis, to determine the historical diversity of strains, and to understand the mechanisms leading to changes in genome structure and virus evolution.

KEYWORDS: virus diversity, phylogenetics, evolution, RNA virus, arbovirus

SUPPLEMENT: Evolution of the Insect Virome

CITATION: Pyke and Warrilow. Archival Collections are Important in the Study of the Biology, Diversity, and Evolution of Arboviruses. *Evolutionary Bioinformatics* 2016:12(S2) 27–30 doi: 10.4137/EBO.S40569.

TYPE: Commentary

RECEIVED: July 13, 2016. **RESUBMITTED:** August 16, 2016. **ACCEPTED FOR PUBLICATION:** August 17, 2016.

ACADEMIC EDITOR: Liuyang Wang, Associate Editor

PEER REVIEW: Five peer reviewers contributed to the peer review report. Reviewers' reports totaled 982 words, excluding any confidential comments to the academic editor.

FUNDING: Authors disclose no external funding sources.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: David.Warrilow@health.qld.gov.au

Viral Taxonomic Methodologies

Virus taxonomy has historically considered multiple contributing types of biological data and hence is referred to as a *polythetic* system of classification (as opposed to *monothetic*).[1] Our understanding of virus evolution has itself changed in parallel with the multiple technological developments that have made this classification system possible, with the emphasis at any one time being on the technology in vogue in that era. The types of data used for the purposes of classification range from what might be considered *absolute* (eg, structural or genome data) to *comparative* (eg, serological). Following the recognition of mosquitoes and ticks as vectors for virus transmission and isolation and propagation of these *arthropod-borne* viruses (arboviruses) in newborn mice, classification was largely dependent on antigenic evidence, which was painstakingly obtained using hemagglutination and various serological techniques.[2] The comparative nature of this analysis meant that it was entirely dependent on sequential viral isolations and the storage of archival serological specimens, both crucial for determining accurate taxonomic groupings and defining the etiology of disease and transmission. Ongoing collections of reference arboviral isolates have indeed provided the backbone for phenotypic, genotypic, and evolutionary studies of these important viruses and enabled significant advances in vaccine and therapeutic agent development. Equally important is the continued development of vector control mechanisms and early diagnosis to assist public health initiatives and mitigation of disease. Thus, it is the intention of this commentary to demonstrate the indisputable importance of arbovirus reference collections and their significance in future research and disease control.

Biosecurity Concerns

The advent of the molecular biology era, with its emphasis on nucleic acids, has led to the suggestion that it might theoretically be possible to develop a mathematically based system of virus classification using only genome sequence data, but it has yet to be demonstrated. On the other hand, it has also been argued that such a system may be practically unfeasible.[3] With increasing emphasis on the storage of genome sequence data rather than the viruses themselves, there is the perception that the value of reference collections has diminished. This is particularly so when the biosecurity risk of carrying such material is considered, particularly in the absence of a vaccine. Between 1980 and 1991, a review of 58 publications revealed that of the 119 reported laboratory-acquired viral infections,

approximately 75% were caused by arboviruses or hantaviruses, and by 1991, the American Committee on Arthropod-Borne Viruses had assigned 535 registered arboviruses to biosafety levels 1–4.[4] In Australia, the Department of Health has designated nonvaccine strains of yellow fever virus as Tier 2 level Security Sensitive Biological Agents (http://www.health.gov.au/ssba).

Archival Collections as Tools in the Battle Against Emerging Arboviruses

Fundamentally, the above considerations potentially apply to all infectious agents. For example, in the case of variola virus, it could be argued that historic stocks of the virus have contributed little to recent research and hence should be destroyed.[5] However, rapid urbanization and greater mobility have increased the epidemic risk of both previously known and new viral pathogens. With the current proportion of emerging viruses such as Ebola, chikungunya virus (CHIKV), and Zika virus (ZIKV), it is clear that we are far from a complete understanding of their unpredictable nature and ability to evolve into new, highly transmissible agents of disease. It is only from methodical analysis of previous strains that we can fully assess the true pathogenic potential of viruses, and this must be continually revised as new and more advanced technological capabilities become available. Therefore, more detailed study of stored viral strains may unlock vital information necessary for future diagnostics and infection control. A good example of this is the use of archival isolates for the development and testing of pan-genus conserved primer sets.[6–8]

Ensuring the Authenticity of Reference Strains

The primary value of reference collections is their inherent comparative capacity such that when new material is collected, it can be directly assessed against agreed reference strains, enabling the classification of material collected worldwide. Reference strains are often the first strain (or prototype) obtained, and their authenticity can only be verified in reference to the original archival specimens. For example, the strain Sofjin was the first isolate of tick-borne encephalitis virus collected and has since become the de facto reference strain. Recent genome sequencing of the Sofjin tick-borne encephalitis virus strain from an original collection was used to determine that a strain that had been designated Sofjin held at another site had most likely been contaminated.[9] This example demonstrates the value of archival material in maintaining the authenticity of reference strains, enabling data in different laboratories to be consistent for the purposes of comparison. Strain authenticity is also imperative in the development of antigens for vaccines and diagnostic platforms. In 1962, it was discovered that a seed lot of the yellow fever virus 17D-attenuated vaccine was contaminated with avian leukosis virus, a potentially oncogenic agent. Rigorous measures had to be undertaken to eliminate the contaminant, and since 1982 only cleared and archived seed stocks of the vaccine have been used.[10]

Preventing Confusion in Nomenclature

Another issue that has arisen in recent years is the repeated sampling of isolates of a given species of arbovirus, and the assignation of multiple names to essentially the same virus. In particular, this has occurred when a virus has been isolated at sites separated by a large geographic area. For example, Malakal virus and Kimberley virus are ephemeroviruses (family *Rhabdoviridae*) which are midge and mosquito-transmitted arboviruses of cattle. These two viruses were collected on separate continents: Malakal virus in 1963 in Africa and Kimberley virus in 1973 in Australia. Intriguingly, they share 90.6% nucleotide identity.[11] Based on current species demarcation for the group, they can be assigned to the same species. Another example is Tibrogargan virus and Bivens Arm viruses, tibroviruses (family *Rhabdoviridae*) isolated in 1976 in Australia and 1981 in the United States, respectively. Similarly, these viruses share 93% nucleotide sequence identity and also should be considered part of the same species.[12] These examples show that it is only by sequencing archival specimens that conflicts regarding accurate species designations can be resolved.

This issue of double naming has been exacerbated when repeated samplings have occurred decades apart. In these cases, one isolate may be characterized using one technology, while the later one is characterized using an entirely different technology. As a result, a connection is never made between the two isolates, and they are designated different names. This was exemplified recently by two orthobunyaviruses (family *Bunyaviridae*): Gan Gan virus and Trubanaman virus, isolated in Australia in 1970 and 1965, respectively. They have both been re-isolated and re-named as either Salt Ash virus (Gan Gan virus) or Murrumbidgee and Buffalo Creek virus (Trubanaman virus), respectively.[13–15] In these cases, the original isolate was characterized using serological techniques, whereas the recent isolations were only characterized by sequencing. Only by sequencing the original isolate was it apparent that the older and newer isolates were in fact the same species. Hence, this example demonstrates the value of maintaining archival material for the purposes of clarifying virus classification in future. This is particularly evident when new technologies emerge which open previously unexplored opportunities for virus characterization.

Responding to Disease Outbreaks

Arboviruses, which are nearly exclusively RNA viruses, are prone to rapid change. New phenotypes can emerge, as was seen recently with the emergence of a mutated, highly transmissible CHIKV strain on Réunion Island and in the Indian Ocean in 2005–2006, resulting in ≈266,000 human cases.[16,17] In 2013, CHIKV was introduced into the Americas following an initial outbreak on the Caribbean Island, Saint Martin.[18]

Surprisingly, phylogenetic analysis demonstrated that the Indian Ocean CHIKV lineage belonging to the East, Central, South African genotype was not responsible for the later epidemics in the Americas, which was caused by a strain belonging to the Asian lineage. In both instances, it was crucial that contemporary circulating CHIKV strains could be compared both genotypically and phenotypically with older African and Asian strains. Similarly, with the advent of ZIKV emergence in the Pacific in 2007 and the Americas in 2015, archival ZIKV strains were essential for comparative analyses. In particular, when infection of pregnant mothers with the 2015 Brazilian strain of ZIKV was associated with cases of congenital malformation in newborns,[19] a rapid cascade of research initiatives were undertaken. Codon usage adaptation has been observed in the ZIKV Asian lineage from which the Brazilian strain was derived.[20] An *in vitro* study comparing the 1947 African strain with the Brazilian strain showed that the latter had significantly different growth properties in human cortical progenitor cells and organoids and induced cell death.[21] It was only by comparing the Brazilian strain with the archival strain that an emerging pathogenic phenotype could be demonstrated.

Understanding Viral Evolution

The rapid change of the RNA viruses, including arboviruses, has resulted in great diversity and adaptability. This does not always translate into viability, as various lineages go to extinction, to eventually be replaced by others. Hence, archival specimens are essential to get an appreciation of the historical diversity of arboviruses. For example, Japanese encephalitis virus (JEV) isolates collected in Indonesia before 1980 were predominantly genotype II (GII), while after that they were replaced with predominantly GIV strains.[22] In northern Vietnam, prior to 1990, the isolates collected were predominantly GIII, while after that date they were replaced with strains of GI.[23] Similarly, in 1995, an incursion of JEV occurred for the first time in the Australasian region in the Torres Strait.[24] The arbovirus reference collection at Queensland Health provides key antigenic material that allowed the first cases to be diagnosed and provided subsequent material for further research. In 2000, a second JEV incursion into the region was identified, and it was discovered that the previously circulating GII virus was replaced in the Torres Strait with a GI virus.[25] Likewise with dengue virus serotype I, an isolate of GII has not been collected since the 1960s, and a subclade of GI has not been collected since the 1940s.[26] It is only the historical collection of these isolates, and the subsequent sequencing when the technology becomes commonplace, which has made these evolutionary studies possible. These examples illustrate the importance of archival specimens to understanding the historical diversity of arboviruses and their evolution.

There are a number of other examples of where archival isolates can shed light on viral evolution. By comparing the archival specimen with newer strains, they can provide direct evidence of genetic restructuring such as recombination or reassortment. For example, reassortment has been documented many times in the family Bunyaviridae.[27] The sequencing of the genome segments of archival specimens of this family will enable the origin of the various reassortant segments to be determined. In addition, archival isolates have valuable sequence information for use in evolutionary analysis. Importantly however, it is only the archival specimens themselves that present an opportunity to compare viral phenotypes at some future time and hence enable the evolution of such traits as virulence to be studied. Finally, by their inclusion in long-term datasets, they can provide the basis for the accurate calculation of the rates of RNA virus evolution.

Conclusions

Given the importance of arbovirus reference collections, a concerted effort should be made to address any biosecurity concerns rather than using this aspect as a means to eliminate them permanently. Archival arbovirus isolates and specimens are an invaluable source of genetic information, enabling researchers to explore a number of important questions on virus evolution and pathogenesis, track the history of disease and assist identification of hosts and reservoirs. Such collections should be adequately maintained to ensure the authenticity of reference material, establish taxonomic relationships, test improvements in diagnostics, and to answer fundamental questions as new technologies emerge.

Acknowledgment

The authors thank Cadhla Firth, Jody Hobson-Peters, and Roy Hall for useful suggestions on the article.

Author Contributions

Wrote the first draft of the manuscript: DW. Contributed to the writing of the manuscript: ATP. Jointly developed the structure and arguments for the paper: DW, ATP. Made critical revisions and approved final version: DW, ATP. Both authors reviewed and approved of the final manuscript.

REFERENCES

1. Condit RC. Principles of virology. In: Knipe DM, Howley PM, eds. *Fields Virology*. Vol 1. 5th ed. Phildelphia, PA: Lippincott Williams & Wilkins; 2007:25–57.
2. Calisher CH, Gould EA. Taxonomy of the virus family Flaviviridae. *Adv Virus Res*. 2003;59:1–19.
3. Simmonds P. Methods for virus classification and the challenge of incorporating metagenomic sequence data. *J Gen Virol*. 2015;96(pt 6):1193–206.
4. Sewell DL. Laboratory-associated infections and biosafety. *Clin Microbiol Rev*. 1995;8(3):389–405.
5. Arita I, Francis D. Is it time to destroy the smallpox virus? *Science*. 2014; 345(6200):1010.
6. Grywna K, Kupfer B, Panning M, et al. Detection of all species of the genus Alphavirus by reverse transcription-PCR with diagnostic sensitivity. *J Clin Microbiol*. 2010;48(9):3386–7.
7. Lambert AJ, Lanciotti RS. Consensus amplification and novel multiplex sequencing method for S segment species identification of 47 viruses of the *Orthobunyavirus*, *Phlebovirus*, and *Nairovirus* genera of the family Bunyaviridae. *J Clin Microbiol*. 2009;47(8):2398–404.

8. Scaramozzino N, Crance JM, Jouan A, DeBriel DA, Stoll F, Garin D. Comparison of flavivirus universal primer pairs and development of a rapid, highly sensitive heminested reverse transcription-PCR assay for detection of flaviviruses targeted to a conserved region of the NS5 gene sequences. *J Clin Microbiol*. 2001;39(5):1922–7.
9. Kovalev SY, Mukhacheva TA, Kokorev VS, Belyaeva IV. Tick-borne encephalitis virus: reference strain Sofjin and problem of its authenticity. *Virus Genes*. 2012;44(2):217–24.
10. Frierson JG. The yellow fever vaccine: a history. *Yale J Biol Med*. 2010; 83(2):77–85.
11. Blasdell KR, Voysey R, Bulach DM, et al. Malakal virus from Africa and Kimberley virus from Australia are geographic variants of a widely distributed ephemerovirus. *Virology*. 2012;433(1):236–44.
12. Lauck M, Yu SQ, Cai Y, et al. Genome sequence of Bivens arm virus, a tibrovirus belonging to the species Tibrogargan virus (Mononegavirales: Rhabdoviridae). *Genome Announc*. 2015;3(2):e89–e15.
13. Gauci PJ, McAllister J, Mitchell IR, et al. Genomic characterisation of three Mapputta group viruses, a serogroup of Australian and Papua New Guinean bunyaviruses associated with human disease. *PLoS One*. 2015;10(1):e0116561.
14. Huang B, Firth C, Watterson D, et al. Genetic characterization of archived bunyaviruses and their potential for emergence in Australia. *Emerg Infect Dis*. 2016;22(5):833–40.
15. Shchetinin AM, Lvov DK, Deriabin PG, et al. Genetic and phylogenetic characterization of Tataguine and Witwatersrand viruses and other Orthobunyaviruses of the Anopheles A, Capim, Guama, Koongol, Mapputta, Tete, and Turlock Serogroups. *Viruses*. 2015;7(11):5987–6008.
16. Chastel C. [Chikungunya virus: its recent spread to the southern Indian Ocean and Reunion Island (2005–2006)]. *Bull Acad Natl Med*. 2005;189:1827–35.
17. Tsetsarkin KA, Vanlandingham DL, McGee CE, Higgs S. A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS Pathog*. 2007;3(12):e201.
18. Cassadou S, Boucau S, Petit-Sinturel M, Huc P, Leparc-Goffart I, Ledrans M. Emergence of chikungunya fever on the French side of Saint Martin island, October to December 2013. *Euro Surveill*. 2014;19(13):20752.
19. de Fatima Vasco Aragao M, van der Linden V, Brainer-Lima AM, et al. Clinical features and neuroimaging (CT and MRI) findings in presumed Zika virus related congenital infection and microcephaly: retrospective case series study. *BMJ*. 2016;353:i1901.
20. de Melo Freire CC, Iamarino A, de Lima Neto DF, Sall AA, de Andrade Zanotto PM. Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans. *BioRxiv*; 2015. Available at: http://biorxiv.org/content/biorxiv/early/2015/11/25/032839.full.pdf.
21. Cugola FR, Fernandes IR, Russo FB, et al. The Brazilian Zika virus strain causes birth defects in experimental models. *Nature*. 2016;534(7606):267–71.
22. Schuh AJ, Guzman H, Tesh RB, Barrett AD. Genetic diversity of Japanese encephalitis virus isolates obtained from the Indonesian archipelago between 1974 and 1987. *Vector Borne Zoonotic Dis*. 2013;13(7):479–88.
23. Do LP, Bui TM, Hasebe F, Morita K, Phan NT. Molecular epidemiology of Japanese encephalitis in northern Vietnam, 1964–2011: genotype replacement. *Virol J*. 2015;12:51.
24. Hanna JN, Ritchie SA, Phillips DA, et al. An outbreak of Japanese encephalitis in the Torres Strait, Australia, 1995. *Med J Aust*. 1996;165(5):256–60.
25. Pyke AT, Williams DT, Nisbet DJ, et al. The appearance of a second genotype of Japanese encephalitis virus in the Australasian region. *Am J Trop Med Hyg*. 2001;65(6):747–53.
26. Pyke AT, Moore PR, Taylor CT, et al. Highly divergent dengue virus type 1 genotype sets a new distance record. *Sci Rep*. 2016;6:22356.
27. Briese T, Calisher CH, Higgs S. Viruses of the family *Bunyaviridae*: are all available isolates reassortants? *Virology*. 2013;446(1–2):207–16.