Arthropods as a source of new RNA viruses

L. Bichaud \textsuperscript{a, b}, X. de Lamballerie \textsuperscript{a, b}, C. Alkan \textsuperscript{a, b}, A. Izri \textsuperscript{a, b}, E.A. Gould \textsuperscript{a, b}, R.N. Charrel \textsuperscript{a, b, *}

\textsuperscript{a} Aix Marseille Universit\é,IRD French Institute of Research for Development,EHESP French School of Public Health,EPV UMR_D 190 "Emergence des Pathologies Virales", 13385 Marseille, France
\textsuperscript{b} IHU Méditerranée Infection,APHM Public Hospitals of Marseille 13385 Marseille, France

\textbf{Article info}

Received 11 August 2014  
Accepted 5 September 2014  
Available online 18 September 2014

\textbf{Abstract}

The discovery and development of methods for isolation, characterisation and taxonomy of viruses represents an important milestone in the study, treatment and control of virus diseases during the 20th century. Indeed, by the late-1950s, it was becoming common belief that most human and veterinary pathogenic viruses had been discovered. However, at that time, knowledge of the impact of improved commercial transportation, urbanisation and deforestation, on disease emergence, was in its infancy. From the late 1960s onwards viruses, such as hepatitis virus (A, B and C) hantavirus, HIV, Marburg virus, Ebola virus and many others began to emerge and it became apparent that the world was changing, at least in terms of virus epidemiology, largely due to the influence of anthropological activities. Subsequently, with the improvement of molecular biotechnologies, for amplification of viral RNA, genome sequencing and proteomic analysis the arsenal of available tools for virus discovery and genetic characterisation opened up new and exciting possibilities for virological discovery. Many recently identified but “unclassified” viruses are now being allocated to existing genera or families based on whole genome sequencing, bioinformatic and phylogenetic analysis. New species, genera and families are also being created following the guidelines of the International Committee for the Taxonomy of Viruses. Many of these newly discovered viruses are vectored by arthropods (arboviruses) and possess an RNA genome. This brief review will focus largely on the discovery of new arthropod-borne viruses.

\textcopyright 2014 Elsevier Ltd. All rights reserved.

\begin{enumerate}
\item \textbf{Introduction}

The 20th century will probably be most well remembered for historical events that include the two World Wars, the development of nuclear weapons, major improvements in transportation, the concept of climate change due to anthropological intervention and the mid-twentieth century, virus discovery was totally reliant on cell and animal culture virus isolation methods, with serological follow up to distinguish individual virus species, genera and strains. The flaviviruses, mosquito-borne yellow fever virus (agent of haemorrhagic fever) and tick-borne Louping ill virus (agent of encephalitis), were the first arboviruses to be discovered \cite{2}. This opened the floodgates for the discovery of more than 500 arboviruses during the next 30 years or more.

More recently, major advances in molecular biotechnological methods, such as the polymerase chain reaction, next-generation sequencing, proteomic analyses, bioinformatics, have augmented the arsenal of tools amenable for virus discovery and further genetic characterization.

Vector-borne infectious diseases cause a significant proportion of the global infectious disease burden. It is estimated that nearly 50\% of the world’s population is infected with at least one type of group of infectious organisms. They were given the name viruses, meaning poison and other noxious substances \cite{1}.

Many of the early discoveries relating to arthropod-borne viruses (arboviruses) resulted from the pioneering studies of scientists associated with the Rockefeller Foundation. From about 1930 until the mid-twentieth century, virus discovery was totally reliant on cell and animal culture virus isolation methods, with serological follow up to distinguish individual virus species, genera and strains. The flaviviruses, mosquito-borne yellow fever virus (agent of haemorrhagic fever) and tick-borne Louping ill virus (agent of encephalitis), were the first arboviruses to be discovered \cite{2}. This opened the floodgates for the discovery of more than 500 arboviruses during the next 30 years or more.

More recently, major advances in molecular biotechnological methods, such as the polymerase chain reaction, next-generation sequencing, proteomic analyses, bioinformatics, have augmented the arsenal of tools amenable for virus discovery and further genetic characterization.

Vector-borne infectious diseases cause a significant proportion of the global infectious disease burden. It is estimated that nearly 50\% of the world’s population is infected with at least one type of
vector-borne pathogen [3,4]. Haematophagous vectors transmit the pathogens between vertebrate hosts when taking a blood meal. The most common vectors of human and animal viral disease are mosquitoes, sandflies, ticks and midges [5]. Importantly, more than two thirds of emerging human pathogens are of zoonotic origin, and of these, more than 70% originate from wildlife species [6,7] in which vector-borne diseases are very well represented [5]. The geographic distribution of the incidence of vector-borne diseases is grossly disproportionate, with the overwhelming impact in developing countries located in tropical and subtropical regions [3]. RNA viruses are notoriously genetically variable, partly due to the high error rate of their RNA polymerases which are the driving force for viral replication and also because many RNA viruses can exchange genetic information by recombination and/or reassortment of viral gene fragments or even entire genes. Consequently, in the sylvatic environment, RNA viruses, are able to exploit this genetic diversity to cross natural species barriers and thereby infect new host species. Consequently, because of this genetic plasticity RNA viruses represent more than 70% of the recognised viruses that infect humans. Thus, the vast majority of emerging viruses of medical interest, as listed by the Institute of Medicine, possess RNA genomes [8].

Several of the recognised major vector-borne human and animal diseases have exhibited dramatic changes in incidence and geographic distribution over the past 3 decades. A variety of factors have contributed to the relative expansion of both geographic and host range; these include: (i) viral adaptation to new vector and host species, (ii) the major influence of commercial transportation, also including global movement of plants and commercial goods, (iii) ecosystem changes that lead to more frequent contact between wildlife and domestic animals or human populations, (iv) human demographics and behaviour, (v) economic development and land use, (vi) international travel and commerce [9]. In addition, climate changes and El Niño/Southern Oscillation (ENSO) may also have contributed to the emergence, expansion of activity area, and epidemics of some vector-borne diseases such as dengue, chikungunya, blue tongue, and Rift Valley fever [10–18].

The aim of this review article is to compile the techniques that have been used in the past and are currently used to look for RNA viruses in arthropods.

2. Why should we look in arthropods for RNA viruses?

Arthropods are invertebrates with jointed limbs, segmented bodies, and exoskeletons made of chitin. They include insects, spiders, crustaceans, and centipedes. Vector-borne diseases are transmitted among their human, animal, or plant hosts by haematophagous (those requiring a blood meal for survival) arthropods, usually insects if they feed on plant hosts, they are not necessarily haematophagous. Because humans and arthropods often share common habitats, arthropods may feed on humans.

Infectious diseases transmitted by arthropods have long been associated with significant human illness and death. From the 17th to early 20th centuries, human morbidity and mortality due to vector-borne diseases outstripped that from all other causes combined [19,20]. Vector-borne diseases have the potential to cause enormous economic harm due to morbidity and mortality of animals, trade embargoes, and vaccination costs [22]; they can disrupt entire ecosystems and, under the right conditions, could potentially expand their host range to include livestock, pets, or humans. These diseases profoundly restrict socioeconomic status and development in countries with the highest rates of infection, many of which are located in the tropics and sub-tropics.

Some experts warn that the arboviruses are “viruses looking for a human disease.”: (i) dengue's resurgence has been marked not only by epidemics, but also by the emergence of a more severe form of disease, dengue haemorrhagic fever [19]; (ii) yellow fever virus, which along with dengue was controlled in the Americas by a variety of mosquito abatement techniques through the mid-20th century, remains a constant threat, for its reintroduction into Aedes-infested metropolitan areas throughout the tropical world [8], and with concerns that it might make its first appearance in Asia [5]; (iii) Rift Valley fever primarily affects livestock but can also be transmitted to humans through direct contact with the tissues or blood of infected animals, as well as by mosquito bites [23,24]. Over the past 30 years dengue fever has re-emerged in Asia and the Americas, West Nile virus (WNV) has spread rapidly throughout the United States, and chikungunya fever has resurfaced in Asia, Africa, the Indian Ocean and now in the Caribbean. Chikungunya fever virus was also introduced into northern Italy where it caused a localised outbreak. It is considered highly likely that this virus will be re-introduced into Europe by visitors returning from the Caribbean [19,25,26]. The world has also recently witnessed the emergence and spread of other vector-borne diseases [27], including bluetongue (a devastating viral disease), transmitted to ruminant livestock by insect vectors, that first appeared in northern Europe in 2006. Several novel phleboviruses transmitted by sandflies and ticks have also emerged in recent decades [28–30]. The considerable economic, ecological, and public health impacts of vector-borne diseases are expected to continue [31]. Therefore, there is a strong case for intensive investigations of arthropods, to establish whether or not many of them are carrying arboviruses with the potential to emerge into human and/or animal populations. One obvious candidate is Japanese encephalitis virus which in Asia has a very wide range of arthropods that can be infected.

3. Available methods of analysing arthropods for the presence of RNA viruses?

3.1. Collection of arthropods from wildlife species

For decades, collecting specimens from wildlife species was the predominant source of zoonotic viruses. During the first half of the 20th century, independent investigators in research and academic centres worldwide, collected field or human clinical samples from which they isolated and characterised the first group A, B and C arboviruses (reviewed in [2]). Subsequently, The Rockefeller Foundation Virus Program was established to investigate arthropod-borne viruses of vertebrates [32] through surveillance of febrile and haemorrhagic diseases. This prestigious, internationally funded programme which was funded from 1951 to 1971, laid the foundation for the discovery of hundreds of arboviruses. In Belem, Brazil, more than 50 new tropical viruses were discovered, including viruses responsible for debilitating, non-fatal jungle fevers common in human populations living in the Amazon region. Kyasanur Forest disease virus was discovered at the Rockefeller laboratory in Pune, India. Crimean haemorrhagic fever virus, discovered in the former Union of Soviet Socialist Republics and Congo disease virus in East Africa were shown to be caused by the same virus and consequently were given a single name, viz., Crimean-Congo haemorrhagic fever virus. Several non-arboviruses related to rabies virus, were discovered in Africa. Rockefeller Virus Programme scientists also discovered viruses in healthy wild animals and arthropods. This innovative approach to disease surveillance identified a number of agents, such as Oropouche virus in Trinidad and Brazil [32], later recognized as the cause of several major tropical epidemics [33]. Most of the current major arboviruses were discovered, characterized, studied, and placed in the International Catalogue of Arboviruses [21] that lists 504 arboviruses...
worldwide, 124 of which have been associated with an animal and/or human disease [34].

During its two decades of existence, the programme was an outstanding catalyst for international surveillance and research. It was the influential work of these centres that largely created the field of arbovirology. The rate of discovery of arboviruses reflected the intensity of the worldwide search, funded by the Rockefeller Foundation, government agencies, and universities from the 1950s to the present time (Table 1). In relatively recent years there has been a significant decrease in activity for such programmes (Fig. 1). However, new arboviruses continue to be discovered, whenever and wherever a search is made. Indeed, the current decade is the first to have shown a renewed and increasing tendency to isolate and describe new arboviruses. Undoubtedly, the recent development and disseminated exploitation of New Generation Sequencing (NGS) which can provide entire genomic sequence data has greatly increased the discovery of arboviruses indeed, several new groups of RNA viruses (both genera and families) remain to be validated by the ICTV. These include the genus Negevirus [35], the family Meso-niviridae [36] the genus Artivirus in the family Tottiviridae [37] and the insect specific flaviviruses for which a new genus is likely to be recommended.

3.2. Which laboratory techniques are available?

3.2.1. Isolation of infectious viruses

Unlike bacteria, viruses require living cells to propagate. This was beginning to be realised during the first decades of the 20th century. Accordingly, scientists experimented with a variety of methods to reproduce human and/or animal viruses in an infectious form to facilitate virus characterisation. These included inoculation of material from infected humans or animals. In the late 1920s and early 1930s successes with yellow fever virus and louping ill virus using mice as the susceptible host proved very successful [2] and enabled virus transmission and immune protection studies. During the 1930s scientists from the Rockefeller Foundation began to experiment with tissues taken from animals and kept alive in culture vessels, being maintained with basic culture medium containing serum proteins. Whilst these methods were proving to be reasonably successful, early successes were observed in tissues derived from the spinal cords of animals. This was not considered a good source for propagation of viruses potentially to be used for human vaccine development. Consequently, Rockefeller scientists experimented with non-neural sources of cells that could support the propagation of poliovirus. By the beginning of the 1950s, methods of cell culture using non-neural animal cells had been developed and these formed the basis for the development of a wide range of animal cell culture based systems that has underpinned human/animal virology for more than 50 years. A major advantage of these methods is that the cells form monolayers on which cytopathic effects can be monitored and used to estimate the level of infectivity (the titre) of each virus under study. Moreover, despite a lull in the popularity of cell culture methods for virus identification, as molecular methods of analysis have developed [38,39], modern cell culture systems are now being increasingly widely used for the isolation and amplification of novel viruses [40–43]. This is largely because it is now being realised that without the demonstration of infectivity and transmissibility, it is not scientifically valid to define a genomic sequence as being that of a novel virus.

3.2.2. Morphological imaging of viruses

The development of the electron microscope (EM) in 1933, provided the first visual images of viruses [44]. Electron microscopy rapidly became a major diagnostic laboratory procedure with which to identify viruses associated with disease. Moreover, description of the morphological characteristics and size of the viruses were established as “founding criteria” for virus taxonomy [45,46]. As the application of the EM to virology became a practical method, rapid diagnostic protocols were developed to enable the visualization of viruses with a few hours of commencement of sample analysis. However, the diagnostic process has the disadvantage that the use of the EM requires substantial operator expertise and a high concentration of virus in the sample under analysis. A further diagnostic refinement is the use of immune-electron microscopy, whereby serum from infected individuals or specific antibodies generated in experimentally virus-infected animal [47] may facilitate more precise identification of the aetiological agent when compared with size and morphological criteria alone. Whilst in the 21st century electron microscopy plays a smaller role than hitherto, in virus research and diagnosis the value of this technology, in competent hands, should not be underestimated.

3.2.3. Detection of viral genomic RNA using PCR amplification

Major developments in virus detection and genomic characterisation have taken place during the past 40 years. The discovery of RNA-dependent DNA polymerase reverse transcriptase in 1970 and the subsequent development of the polymerase chain reaction (PCR) together with the realisation that viral RNA polymerases in general have low fidelity has considerably modified the approach to screening arthropods for viruses. Multiplex (multiple-sample studies) PCR assays in current application can detect and identify
genomic products using fluorescent tags, barcodes, mass spectrometry and capillary electrophoresis [48–52]. Even higher levels of multiplexing can be obtained with DNA microarrays [53,54]. Although, sensitivity is insufficient for clinical application it is considered sufficiently sensitive for virus discovery from arthropods. Ultimately, the sensitivity is likely to be increased through the development of diagnostic tools based on the introduction of nanofluidics and electronic nanocircuits [55–57]. For a comprehensive list and detailed description of molecular methods currently being used for virus discovery, readers are referred to the following review [42].

3.2.4. Next-generation sequencing

Novel DNA sequencing techniques based on sequence-independent amplification and subsequent sequence determination, known as “Next-Generation Sequencing” (NGS) techniques, are new tools that can provide high-throughput sequence data of any viral genomes in a given sample without the need for previous knowledge about their nature. Thus, these new techniques can reveal limitless insight about the genome, transcriptome and epigenome of any species. It is important to underline that the aim of this study is not to describe the advantages and pitfalls of the different techniques currently available for NGS, i.e. 6 at present time: Roche, SOLID, Illumina, Helicos, PacBio, and Ion Torrent. It is rather to present how these recent techniques can be combined with conventional techniques. Such platforms allow the discovery of novel viral species within a complex sample. Due to the decreasing costs, their usage is increasing exponentially.

The advent of these rapid platforms for RNA/DNA sequencing has revolutionized microbiology, by providing culture-independent methods amenable to high throughput sample screening for virus discovery. Over the past 10 years, the per base cost of sequencing has decreased 10,000-fold, from $5000 per megabase using capillary electrophoresis to $0.5 per megabase using the Illumina platform [58,59]. The techniques are already being routinely used in thousands of laboratories. Accordingly, the volume of sequence data has also grown exponentially. Irrespective of the platform chosen, the process following sequence acquisition is similar for pathogen discovery: contiguous sequences are assembled, host sequence is eliminated, and the residual sequences are examined for similarity to known microbial sequences at the nucleotide or amino acid level [44,60]. Almost overnight, the opportunities for novel virus discovery has increased beyond the wildest dreams of virologists less than 20 years ago.

4. What is already being discovered about viruses in arthropods?

The first relatively wide survey of viral diversity within mosquitoes using metagenomics was published in 2011 [61]. More recently, insect-specific viruses belonging to the Flaviviridae and Birnaviridae families were detected in mosquitoes in southern France [62] and the use of metagenomics led to the discovery of a large number of known and unknown insect-specific or zoonotic viruses associated with arthropods [63,64]. Most recently, in a metagenomic analysis of Australian mosquitoes the authors stated that “arthropods can be considered as goldmines for virus discovery programs” [65]: in this study, viruses belonging to several families were identified: Edge Hill virus (Flaviviridae), Wallal virus (Reoviridae), Ross River virus (Togaviridae) and also new viruses such as a dipteran-mammal-associated rhabdovirus called dimarhabdovirus, and sequences belonging to new bunyaviruses. However, this type of broad survey of viral diversity in arthropods is only in its infancy. Nevertheless, metagenomics is a powerful tool for equivalent future studies and will undoubtedly lead to the discovery of many more novel viruses.

An additional incentive for the utilisation of arthropods arises from the ease with which they can be collected and analysed. For example, arthropods may now be used as sentinels for the analysis and prediction of likely emerging arbovirus diseases via surveillance programs that focus on human and veterinary health. This was exemplified in 2011 when an unknown outbreak of fever and thrombocytopenia involving hundreds of patients occurred in rural China [30,66]. Unbiased NGS analysis of pooled patient serum samples was used to identify the causal agent as a novel, highly divergent bunyavirus in the Phlebovirus genus. Furthermore NGS has been used to enable whole-genome sequencing and assembly of highly divergent viruses identified from unknown cultures exhibiting cytopathic effect in cell culture. Heartland virus, a presumed novel tick-borne bunyavirus in the genus Phlebovirus associated with 2 cases of severe febrile illness in hospitalized patients in Missouri [29], and Lone Star virus, another phlebovirus infecting the Amblyomma americanum tick [67], were both successfully sequenced from virally infected cell culture supernatant using NGS.

5. Conclusions

This review has briefly addressed the history of arbovirus detection, identification (including taxonomic issues), characterisation and molecular analysis. Prior to the development of molecular
biological technology that provides precise genetic, proteomic and bioinformatics data, virologists had depended on a variety of physical and biological procedures that included, virus size estimation (usually via filtration methods), morphological characteristics (via electron microscopy), virus isolation and amplification methods, (via tissue culture and/or animal infection), immunological methods (for specific virus identification). Human and animal virology, as well as viruses that infect all other living species, has now reached the position where it is possible to identify whole viral genomic sequences in field material without demonstrating, by culture in susceptible cells, that these genomes represent infectious transmissible viruses. Thus, virology is now entering a new era in which it can conditionally identify hundreds or more likely, thousands of novel viruses but it must decide whether or not a sequence that appears to correspond to that of a novel virus but which cannot be shown to be infectious is eligible for inclusion in the ICTV catalogue. This is a new and exciting time in virology. We must not miss the opportunity to exploit the new technologies but as more potential viral sequences are identified, we must also consider carefully the true meaning of the words “virus species”.

Acknowledgements

This work was supported in part by (i) the project PREDEMICS FP7/2007–2013 n° 278433 (http://predemics.biomedtrain.eu/cms/), (ii) the European Virus Archive (European Union Seventh Framework Programme under grant agreement no. 228292) (http://www.european-virus-archive.com/); (iii) the EDENext FP7-n° 261504 EU project and this paper is catalogued by the EDENext Steering Committee as EDENext#262 (http://www.edenext.eu). The work of RNC was done under the frame of EurNegVirus COST Action TD1303.

References


