



GENOME WATCH

Bugs full of viruses

Astrid Gall

This month's Genome Watch highlights how high-throughput sequencing has provided new insights into the diversity, evolution and genome organization of arthropod viruses.

Arthropods form the largest phylum in the animal kingdom and comprise a remarkable diversity of species. They actively interact with a wide range of other organisms including fungi, plants and vertebrates, and they can therefore act as both a source and a sink for viruses in an ecosystem. Some arthropods, such as mosquitoes or ticks, are important vectors for well-studied viral diseases, but the overall genetic diversity and evolution of arthropod viruses has remained largely unexplored. Three recent studies have used high-throughput sequencing to gain new insights into the diversity, evolution and genome organization of arthropod viruses.

Li *et al.*¹ surveyed negative-sense RNA viruses harboured by arthropods in China. A total of 70 host species were assessed, including mosquitoes, flies, cockroaches, water striders, ticks, spiders, shrimps, crabs and millipedes. RNA sequencing, *de novo* assembly and similarity searches identified 112 distinct RNA-dependent RNA polymerase (*RdRp*) sequences of novel negative-sense RNA viruses, with evidence for at least 16 potentially new families. Phylogenetic analysis revealed that most of the newly described arthropod viruses were basal to known major virus groups, indicating that arthropod viruses may be evolutionary ancestors of viruses that cause disease in vertebrates and plants. Some new lineages filled in 'phylogenetic gaps' between existing viral

genera or families. Importantly, the putative new family *Chuviridae* — named after reaches of the Yanqzi river — fell between the major groups of unsegmented and segmented viruses, and included viruses with various genome organizations, ranging from unsegmented to bi-segmented and circular genomes.

This study also challenges our understanding of the biology of *Bunyaviridae*, a family that includes important vector-borne RNA viruses. Previously described bunyavirus genomes consist of a small (S), medium (M) and large (L) segment, which encode the nucleocapsid, glycoproteins and RdRp, respectively. Similarity searches for viral proteins other than RdRp found segments of the newly identified bunyavirus genomes that encoded structural proteins. However, for some of these new viruses, including a group of tick phleboviruses, the M segments seemed to be missing. Although it cannot be excluded that the apparent absence of M segments is an experimental artefact, these results challenge the concept of what represents a complete virus genome. Tokarz *et al.*² reported another four new bunyavirus genomes lacking M segments in human-biting tick species from the United States. This study also further expanded the known repertoire of arthropod viruses by identifying sequences from one new *Mononegavirales*-like virus and three new positive-sense single-stranded RNA viruses with similarity to invertebrate and plant viruses.

The circular genome structure reported for the *Chuviridae* by Li *et al.*¹ is not commonly found in negative-sense RNA viruses. Circular genomes have previously been described in DNA viruses, and these can be recovered by a combination of rolling circle amplification and high-throughput sequencing. Using this approach, Dayaram *et al.*³ identified 31 new CRESS (circular replication-associated protein encoding single-stranded) DNA viruses in dragonflies and damselflies from the

United States. Seven of these had similarities to the proposed gemcircularvirus group. The host species (*Sclerotinia sclerotiorum*, a plant fungal pathogen) was known for only one of the viruses, and others were similar to viruses associated with animal faecal matter. This study demonstrated that analysis of top-end predators of an ecosystem, such as dragonflies, is an efficient strategy for virus discovery, as viruses (including plant viruses) accumulate in these species.

The largely unbiased high-throughput sequencing approaches used in these studies provide valuable insights both into the diversity of arthropod viruses and into general aspects of viral evolution, but important limitations also need to be considered. The specific protocols used for sample processing and high-throughput sequencing can have a major influence on the genome types, and hence virus families, that are discovered. Furthermore, although many partial or complete virus genomes were identified, virus isolation was either not attempted or was unsuccessful². Finally, additional studies will be necessary to determine whether these arthropod viruses cause diseases in other species and to assess the benefits of including arthropods in surveillance programmes as easily accessible sentinels.

Astrid Gall is at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

e-mail: microbes@sanger.ac.uk

doi:10.1038/nrmicro3479

Published online 13 April 2015

1. Li, C. X. *et al.* Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* **4**, e05378 (2015).
2. Tokarz, R. *et al.* Virome analysis of *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks reveals novel highly divergent vertebrate and invertebrate viruses. *J. Virol.* **88**, 11480–11492 (2014).
3. Dayaram, A. *et al.* Identification of diverse circular single-stranded DNA viruses in adult dragonflies and damselflies (Insecta: Odonata) of Arizona and Oklahoma, USA. *Infect. Genet. Evol.* **30**, 278–287 (2015).

Competing interests statement

The author declares no competing financial interests.

