Virology ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at SciVerse ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus

Kate S. Baker [a,b], Richard M. Leggett [c], Nicholas H. Bexfield [a], Mark Alston [c], Gordon Daly [a], Shawn Todd [d], Mary Tachedjian [d], Clare E.G. Holmes [d], Sandra Crameri [d], Lin-Fa Wang [d], Jonathan L. Heeney [a], Richard Suu-Ire [e,f], Paul Kellam [g], Andrew A. Cunningham [b], James L.N. Wood [a], Mario Caccamo [c], Pablo R. Murcia [h,*]

[a] University of Cambridge, Department of Veterinary Medicine, Madingley Rd, Cambridge, Cambridgeshire, CB3 0ES, United Kingdom
[b] Institute of Zoology, Zoological Society of London, Regent's Park, NW1 4RY, United Kingdom
[c] The Genome Analysis Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom
[d] CSIRO Australian Animal Health Laboratories, Portarlington Road, East Geelong, VIC 3219, Australia
[e] Wildlife Division of the Forestry Commission, Accra, Ghana
[f] University of Ghana, Noguchi Memorial Institute for Medical Research, Legon, Accra, Ghana
[g] Wellcome Trust Sanger Institute, A1301, Hinxton, Cambridgeshire CB101, United Kingdom
[h] MRC—University of Glasgow Centre for Virus Research, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G61 1QH, United Kingdom

ARTICLE INFO

Article history:
Received 19 January 2013
Returned to author for revisions
21 February 2013
Accepted 11 March 2013

Keywords:
Virome
Bat
Megabat
Poxvirus
Viral emergence
Metagenomics
Adenovirus

ABSTRACT

Viral emergence as a result of zoonotic transmission constitutes a continuous public health threat. Emerging viruses such as SARS coronavirus, hantaviruses and henipaviruses have wildlife reservoirs. Characterising the viruses of candidate reservoir species in geographical hot spots for viral emergence is a sensible approach to develop tools to predict, prevent, or contain emergence events. Here, we explore the viruses of *Eidolon helvum*, an Old World fruit bat species widely distributed in Africa that lives in close proximity to humans. We identified a great abundance and diversity of novel herpes and papillomaviruses, described the isolation of a novel adenovirus, and detected, for the first time, sequences of a chiropteran poxvirus closely related with *Molluscum contagiosum*. In sum, *E. helvum* display a wide variety of mammalian viruses, some of them genetically similar to known human pathogens, highlighting the possibility of zoonotic transmission.

Introduction

Zoonoses caused by unknown agents represent a significant proportion of the challenge of emerging infectious diseases (EIDs) (Morens et al., 2004). Viruses account for approximately 25–44% of all EIDs (Jones et al., 2008; Taylor et al., 2001) and studies suggest they are the pathogen class most likely to emerge (Cleaveland et al., 2007; Dobson and Foufopoulos, 2001). Hantaviruses, henipaviruses, SARS coronaviruses and filoviruses are all viruses of zoonotic origin. Nearly 80% of zoonotic EIDs originate from wildlife, and the overall contribution of wildlife pathogens to human EID events is increasing and represent an ongoing threat to global health (Cleaveland et al., 2007; Jones et al., 2008). For example,
a novel coronavirus associated with acute respiratory disease was recently diagnosed in pneumonia patients in Saudi Arabia and London (Bermingham et al., 2012; Zaki et al., 2012). Analysis of the novel coronavirus genome suggests a possible bat origin (Bermingham et al., 2012).

South and East Asia, Eastern Europe, Latin America and tropical Africa constitute areas of increased relative risk for zoonotic emergence from wildlife (Jones et al., 2008; Morens et al., 2004). Numerous studies have successfully combined metagenomics with next generation sequencing to explore the viruses of different animal species, including: domestic pigs and turkeys (Day et al., 2010; Shan et al., 2011); Californian sea lions (Li et al., 2011); and

* Corresponding author. Fax: +44 141 330 2271.
E-mail addresses: kf281@cam.ac.uk (K.S. Baker), richard.leggett@tgac.ac.uk (R.M. Leggett), nb289@cam.ac.uk (N.H. Bexfield), mark.alston@bbsrc.ac.uk (M. Alston), gd311@cam.ac.uk (G. Daly), shawn.todd@csiro.au (S. Todd), mary.tachedjian@csiro.au (M. Tachedjian), clare.holmes@csiro.au (C.E.G. Holmes), sandra.crameri@csiro.au (S. Crameri), linfa.wang@csiro.au (L.-F. Wang), jlh66@cam.ac.uk (J.L. Heeney), suuire@hotmail.com (R. Suu-Ire), pk5@sanger.ac.uk (P. Kellam), a.cunningham@ioz.ac.uk (A.A. Cunningham), jlnw2@cam.ac.uk (J.L.N. Wood), mario.caccamo@tgac.ac.uk (M. Caccamo), pablo.Murcia@glasgow.ac.uk (P.R. Murcia).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.03.014

rodents (Phan et al., 2011). Characterising the viruses of candidate reservoir species in high-risk geographical areas is an important step toward better understanding viral emergence.

Bats are the primary reservoirs for many viral zoonoses, including henipaviruses, filoviruses, some lyssaviruses and SARS-like coronaviruses (Halpin et al., 2000; Kuzmin et al., 2008; Li et al., 2005; Luby et al., 2009; Towner et al., 2007). Indeed, seminal work has been recently published on the role of bats as natural reservoirs of paramyxoviruses (Drexler et al., 2012). Detailed studies of the viruses of insectivorous bat species in both North America and China have been conducted (Donaldson et al., 2010; Ge et al., 2012; Li et al., 2010a; Wu et al., 2012). These studies found large numbers of insect and plant viruses, which were thought to reflect dietary inputs, as well phage sequences and mammalian viruses. The majority of the mammalian viruses identified in those studies were those previously identified in bats (often with high diversity being reported in individual populations) and include: *Adenoviridae* (Li et al., 2010c); *Parvoviridae* (Li et al., 2010b); *Circoviridae* (Ge et al., 2011); *Coronaviridae* (Tang et al., 2006; Woo et al., 2006) and *Astroviridae* (Chu et al., 2008). *Papillomaviridae* and *Herpesvirdae* sequences were also commonly found (Donaldson et al., 2010; Ge et al., 2012; Wu et al., 2012) and some studies also reported *Picornaviridae*, *Flaviviridae* and *Retroviridae* (Li et al., 2010a; Wu et al., 2012).

If we consider that the ~1200 bat species constitute approximately 20% of the class *Mammalia* and that they are near-globally distributed, the benefits of expanding our knowledge of bat viruses on geographic and taxonomic levels become evident. Here we conducted a metagenomic study to detect viruses of *E. helvum*, a frugivorous African bat species that is widely-distributed and migratory throughout much of sub-Saharan Africa. The species is eaten as bushmeat, and the populations studied have ample opportunities for human contact, including a roost directly over a hospital in Accra, Ghana (Hayman et al., 2012).

Results

Bioinformatics analysis

Performance comparison among assemblers

Here, we show important performance differences among four different assemblers (Velvet, ABySS, MetaIDBA and MetaCortex) and three sample types.

The assemblers generated different numbers of contigs, with MetaCortex and ABySS producing more sequences than Velvet and MetaIDBA for each sample type (Fig. 1A). Sample type also affected the number of contigs, with increasing cellularity resulting in more contigs (Table 1), except with MetaIDBA, where the sample allowed to iterate over a kmer size-range generated the most contigs. Contig length parameters also varied with assembler and sample type. Velvet generated contigs with the longest average length across sample types, while ABySS typically produced the longest contigs. Regarding sample type, the longest contigs were generated from the throat sample for each assembler (Table 1).

As well as generating contigs of differing length and number, the nucleotide composition of contigs varied among assemblers. Base composition of the total assembled contigs revealed important differences (Supp. Fig. 2A). Velvet and MetaCortex contigs had similar base compositions, while ABySS contigs incorporated non-ATCG notations (e.g. N, R, Y, Supp. Fig. 2A), and MetaIDBA contigs were primarily composed of adenine (though this was not true for MetaIDBA contigs included in final analyses (Supp. Fig. 2B)).

Consolidation of contigs combines the strengths of assembler approaches and reduces complexity

De novo assemblers have different contig-construction methods, resulting in strengths and weaknesses in different situations. By consolidating contigs from multiple assemblers, we combined the strengths, while reducing the computational complexity of analyzing contigs from assemblers separately. The proportion of contigs retained after consolidation differed among assemblers. For example, ≤22% of ABySS contigs but ≥94% of MetaIDBA contigs were retained into the final consolidated set for each sample type (Table 2). The consolidation resulted in the discard of approximately 30% of the total assembled contigs (~4.4 million sequences). Consistent with the observed assembler differences in contig generation, a variable proportion of sequences were also length-excluded per assembler, but these proportions were approximately equal among sample types (Fig. 1B, Table 2).

Eidolon helvum samples contain large numbers of viral sequences

To identify viral sequences in the consolidated contigs, we used multiple algorithms, which had different efficacies. While BLASTn identified 258 suspect-viral sequences among the sample types, BLASTx and tBLASTx identified a further 6448 and 2563 viral sequences, respectively. Manual exclusion and curation of these 9269 suspect-viral sequences was used to further focus analysis on viral sequences of interest (Fig. 1B). Here, we aimed to explore the viruses that likely infect *E. helvum* (for which they probably constitute a natural reservoir) not the viruses infecting their dietary inputs or bacterial flora. Consequently, 8095 suspect-viral sequences related to viral families not known to infect vertebrates were excluded from further analysis. Subsequent to close inspection of the remaining 1174 suspect-viral sequences, a further 11 sequences were removed due to incorrect classification in the database (not shown). This resulted in 1363 viral sequences related to eight mammalian-infecting viral families being identified. While the majority (77%) were related to viruses with double stranded DNA (dsDNA) genomes, 21% were related to retroviruses (classified separately as sequences may have derived from exogenous-RNA or proviral-DNA forms) and single-stranded DNA and positive-sense single-stranded RNA viruses were also present (Table 3, Fig. 1C).

All sample types, assembly algorithms and BLAST comparison algorithms identified viral sequences (Table 3, Fig. 1C). By using multiple assembly and identification algorithms we generated more contigs and identified more viral sequences.

Analysis of viral sequences by family

Herpesviridae

We identified 539 sequences related to herpesviruses, mostly from the throat sample(Table 3). Sequences related to a wide range of genes and proteins involved in diverse functions including gene regulation, nucleotide metabolism, DNA replication as well as envelope glycoproteins and other structural proteins. Most sequences related to members of the *betaherpesvirinae* ($n=366$) and *gammaherpesvirinae* ($n=171$), and only two sequences most closely related with *alphaherpesvirinae* (Supp. Table 1). Phylogenetic analysis of a region of the DNA polymerase showed the presence of distinct herpesviruses in the throat sample, including some related to other bat betaherpesviruses (Zhang et al., 2012) and a novel gammaherpesvirus (Fig. 2). The presence of contig th_687866 in the throat sample was confirmed by PCR and sequencing (not shown).

Please cite this article as: Baker, K.S., et al., Metagenomic study of the viruses of African straw-coloured fruit bats: Detection of a chiropteran poxvirus and isolation of a novel adenovirus. Virology (2013), http://dx.doi.org/10.1016/j.virol.2013.03.014