Infection, Genetics and Evolution 45 (2016) 242–245

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

Complete genome characterisation and phylogenetic position of Tigray hantavirus from the Ethiopian white-footed mouse, *Stenocephalemys albipes*

Joëlle Goüy de Bellocq [a,b,*], Jana Těšíková [a,c], Yonas Meheretu [d], Dagmar Čížková [a], Anna Bryjová [a], Herwig Leirs [b], Josef Bryja [a,c]

[a] *Institute of Vertebrate Biology of the Czech Academy of Sciences, Brno, Czech Republic*
[b] *University of Antwerp, Evolutionary Ecology Group, Antwerp, Belgium*
[c] *Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic*
[d] *Mekelle University, Department of Biology, Mekelle, Ethiopia*

ARTICLE INFO

Article history:
Received 31 May 2016
Received in revised form 8 August 2016
Accepted 8 September 2016
Available online 09 September 2016

Keywords:
Hantavirus
Murinae
Ethiopia
High throughput sequencing
Genomics

ABSTRACT

Hantaviruses, well-known human pathogens, have only recently been identified on the African continent. Tigray virus (TIGV) was found in Ethiopia in 2012 in a Murinae species, *Stenocephalemys albipes*, but the genetic data obtained at that time were too limited to correctly assess its phylogenetic position within the hantavirus tree. We used high throughput sequencing to determine the complete genome of TIGV, which showed a typical hantavirus organisation. The large (L), medium (M), and small (S) genome segments were found to be 6532, 3594 and 1908 nucleotides long, respectively, and the 5′ and 3′ termini for all three segments were predicted to form the panhandle-like structure typical for bunyaviruses. Nucleotide-based phylogenetic analyses revealed that all three coding segments cluster in the phylogroup III sensu Guo et al. (2013). However, while TIGV S segment is basal to the Murinae-associated hantaviruses, the M and L segments are basal to the Soricomorpha-associated hantaviruses. TIGV is the first Murinae-borne hantavirus showing this inconsistent segmental clustering in the hantavirus phylogenetic tree. We finally propose non-exclusive scenarios that could explain the original phylogenetic position of TIGV.

© 2016 Published by Elsevier B.V.

Hantaviruses (genus *Hantavirus*, family Bunyaviridae) are RNA viruses carried by rodents, soricomorphs (shrews and moles) and bats. Some of these viruses can cause serious health issues in humans; they are the etiologic agents of haemorrhagic fever with renal syndrome and hantavirus cardiopulmonary syndrome. The evolutionary history of this group is complex with evidence of segmental reassortment events, cross-species transmission and virus-host co-divergence (Bennett et al., 2014; Guo et al., 2013).

Hantaviruses were first described from the African continent relatively recently. Initially Sangassou (SANGV), from the murine species *Hylomyscus simus*, was found in Guinea in 2006 (Klempa et al., 2006). Since then, African hantaviruses have been reported in bats (Sumibcay et al., 2012; Weiss et al., 2012; Witkowski et al., 2016) and shrews (Gu et al., 2013; Kang et al., 2011; Kang et al., 2014; Klempa et al., 2007). During a previous study screening for hantaviruses and arenaviruses in Ethiopian small mammals, we found that the Ethiopian White-footed mouse, *Stenocephalemys albipes*, harboured a new hantavirus that we called Tigray (TIGV) (Meheretu et al., 2012). To this point SANGV and TIGV are the only indigenous Murinae-borne hantaviruses found in Africa. The initial TIGV analysis (based on a small part of the L gene) did not allow the phylogenetic relationship of this new Murinae-borne virus with the other hantaviruses to be resolved (Meheretu et al., 2012). Here we sequenced the full genome of TIGV by high throughput sequencing and analysed its phylogenetic position within the hantavirus tree.

Since no fresh or frozen tissue sample was available for TIGV isolation attempts on cell culture we chose one positive sample from the previous study (Meheretu et al., 2012) for next generation sequencing (ET2121; Golgolnaele; 13°52′N, 39°43′E, elevation 2700 m). This sample was positive for Tigray strain 92 (Meheretu et al., 2012). We used a piece of liver preserved in RNAlater reagent (Qiagen) and stored at −80 °C. The RNA was extracted following the viral enrichment protocol S3 described in Dupinay et al. (2014) with 2 modifications: first the tissue (about 50 mg) was homogenised in 500 μL of 1 × PBS buffer instead of HBSS before going through 3 cycles of freezing and thawing; the second modification was that the RNA digestion was performed using 25 U *RNase* ONE *Ribonuclease (Promega)* and 30 U of *Benzonase (Novagen)* at 37 °C for 75 min. Viral encapsidated RNA and residual host nucleic acids

* Corresponding author at: Research Facility Studenec, Institute of Vertebrate Biology of the Czech Academy of Sciences, Studenec 122, 67502 Konesin, Czech Republic.
E-mail address: joellegouy@gmail.com (J. Goüy de Bellocq).

http://dx.doi.org/10.1016/j.meegid.2016.09.009
1567-1348/© 2016 Published by Elsevier B.V.

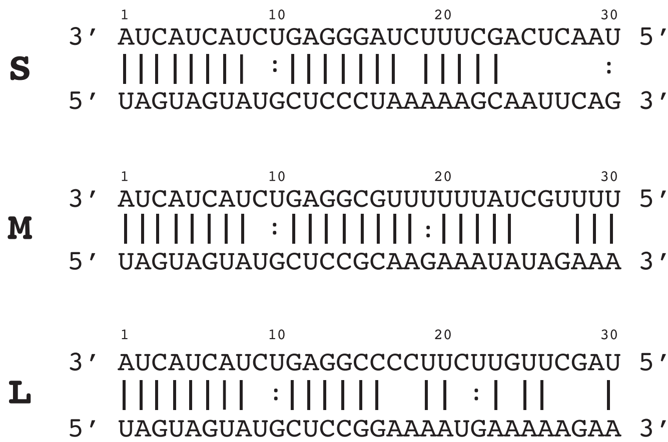

Fig. 1. Panhandle-forming terminal nucleotides of each genomic RNA segment of Tigray (TIGV): small (S), medium (M) and large (L). Vertical lines reflect base pairs and colons the non-canonical U-G pairing.

were then extracted using the RNeasy mini kit (Qiagen). We performed two independent extractions that were finally pooled and eluted in a final volume of 17 μL using RNAeasy MinElute Cleanup kit (Qiagen). Although the amount of RNA at this stage was too low to be detectable by the Qubit RNA HS Assay Kit (Life Technologies), TIGV RNA was detectable by One-step-RT-PCR (SuperScript One-Step RT-PCR System, Invitrogen) using primers from Klempa et al. (2006). We outsourced the library preparation and the sequencing to ViroScan3D (Lyon, France; www.viroscan3d.com). At the ViroScan3D facility the RNA sample was again quantified using the QuantiFluor RNA System (Promega) and 2100 Bioanalyzer (Agilent Technologies) but again no RNA was detectable. The Ovation RNA-Seq system V2 (NuGEN) was used for cDNA generation before library preparation with the Ovation Ultralow Library System (NuGEN). The library was paired-end (PE) sequenced using a 100 bp protocol with indexing on a HiSeq 2500 (Illumina) sequencer in pools of 7 specimens per lane on two lanes (the 6 other specimens were arenaviruses). Reads were processed using Bcl2fastq 2.16.0.10 and demultiplexed based on index sequences. In total we obtained 39,641,687 PE reads for the TIGV sample. FastQC was used for quality checking.

Reads were analysed in Geneious 9 (Kearse et al., 2012). We first trimmed the regions with more than a 5% chance of error per base. A first BLAST of 100 reads showed that about 10% of the reads belong to TIGV. Since this first figure predicted a high number of viral reads in our data, we used only a subset of the reads (5,000,000) to reconstruct the TIGV genome, making the computation time tractable. First we performed a de-novo assembly using Geneious Assembler with the lowest sensitivity parameter and with all other parameters set to default. The 25 first contigs were BLASTed and parts of the contigs with high similarities with M and L hantavirus segments saved to be used as TIGV reference sequences for a subsequent mapping of the reads. We used 6 partial L sequences of about 400 nucleotides (nt) and 3 partial M sequences of about 250 nt as TIGV reference sequences in Geneious Mapper to step-by-step reconstruct the TIGV genome, each iteration adding 80 nt at each end of the reference sequences. The sensitivity was set to low and minimum mapping quality to 30. To reconstruct the S segment, we proceed as for the M and L segments but we used as a reference sequence a sequence of 385 nt that we previously generated by Sanger sequencing using primers from Arai et al. (2008). After completing the assembly of the full TIGV genome, we assessed the total number of reads in the whole dataset that belong to TIGV by using Geneious Mapper. In total 11.7% reads mapped to TIGV (1.8%, 9.1% and 0.8% mapping to the L, M and S segments, respectively) with a mean base coverage of 22,156 (minimum: 2734; maximum: 58,156) for the L segment, 208,692 (min: 3600; max: 522,605) for the M segment and 41,607 (min: 1032; max: 191,155) for the S segment.

The TIGV complete L segment sequence was found to be 6532 nucleotides (nt) long and to contain a single open reading frame (ORF) of 6453 nt (positions 38 to 6493) that encodes the 2151 amino-acid (aa)-long RNA-dependent RNA polymerase (L protein). The complete M segment consists of 3594 nt and again carries a single ORF (3408 nt; positions 41 to 3451) of the 1136 aa-long glycoprotein precursor (GPC). Finally the S segment was found to be 1908 nt long and to contain a single ORF of 1284 nt (positions 47 to 1333) that encodes the 428 aa-long nucleocapsid (N) protein. The usual conserved regions functionally relevant in hantaviruses were also present in TIGV such as the RNA binding domain of the N protein (aa positions 172 to 214) (Xu et al., 2002) and the pentapeptide motif WAASA determining the cleavage of the GPC into the glycoproteins G1 and G2 (aa positions 645 to 649) (Löber et al., 2001). Finally the 3′ and 5′ termini of all three segments of TIGV were typical of bunyaviruses: they were predicted to form panhandle-like structures with the incomplete complementarity at position 9 and the noncanonical U-G pair at position 10 (Fig. 1) which is regularly found in other hantaviruses (Plyusnin et al., 1996). The sequences determined in this study were deposited in GenBank (accession numbers KU934008–KU934010).

Coding parts of TIGV nucleotide sequences were aligned with representatives of the main hantavirus clades (see Supplementary Table 1) at the amino-acid level in Geneious using MUSCLE (Edgar, 2004). We used MEGA 6.06 (Tamura et al., 2013) to evaluate the fit of 24 nested models of nucleotide substitution to the sequences by using the Bayesian information criterion (BIC). The BIC indicated that the substitution model best fitting the data for each of the 3 segments was the GTR $+$ I $+$ Γ. Phylogenetic analyses were performed using maximum likelihood (ML) using PhyML 3.1 (Guindon et al., 2010) and Bayesian inference implemented in MrBayes 3.2.2 (Ronquist et al., 2012). For the ML tree, support was evaluated by 1000 replicate bootstraps. In MrBayes, we used the default priors for all parameters and two independent runs were conducted with 1,000,000 generations per run; trees and parameters were sampled every 500 generations. Runs were initiated from random trees, and three hot chains plus one cold chain were used in all analyses. Convergence was assessed by examining the average standard deviation of split frequencies and the potential scale reduction factor. For each run, the first 25% of trees sampled were discarded as burn-in. Bayesian posterior probabilities were used to assess branch support. Trees were annotated in FigTree, version 1.4.1. (http://tree.bio.ed.ac.uk/software/figtree/).

ML and Bayesian phylogeny estimation produced relatively similar trees (Fig. 2) (the minor differences are visible at nodes with missing bootstrap support). For all 3 genomic segments, TIGV was found within the phylogroup III (sensu Guo et al., 2013) grouping Murinae- and Soricomorpha-borne hantaviruses (Fig. 2). However, TIGV clustering within this clade differed with respect to its genomic segment: for the S segment, TIGV was well supported sister lineage to the clade grouping all Murinae-borne hantaviruses (posterior probability 1, ML bootstrap 89); however for the M and L segments TIGV was sister to the Soricomorpha-borne hantaviruses, even if at L segment this relationship being weakly supported (1/91 and 0.98/61 support for M and L, respectively). This finding is currently unique for a Murinae-borne hantaviruses: all previous cases consistently formed a strongly supported monophyletic cluster within phylogroup III. The unique position of TIGV may be explained by different non-exclusive scenarios: i) historical segment re-assortment between the ancestors of Murinae and Soricomorpha hantaviruses. This scenario seems likely as the wider evolutionary history of hantaviruses shows evidence for both segment re-assortment and host switching (Bennett et al., 2014; Guo et al., 2013); ii) Ethiopian highlands form the most important African centre of endemism, with several endemic mammalian genera (Yalden and Largen, 1992), including the rodent genus *Stenocephalemys*, the only known host of TIGV. These rodents (as well as many other taxa) probably evolved in Ethiopia in complete separation since late Miocene (e.g. Lecompte et al., 2008) and if this is also the case for hantaviruses, we

Download English Version:

https://daneshyari.com/en/article/2822882

Download Persian Version:

https://daneshyari.com/article/2822882

Daneshyari.com