Viral metagenomics demonstrates that domestic pigs are a potential reservoir for Ndumu virus

Masembe *et al.*

VIROLOGY JOURNAL

SHORT REPORT   **Open Access**

Viral metagenomics demonstrates that domestic pigs are a potential reservoir for Ndumu virus

Charles Masembe[1*], George Michuki[2], Maria Onyango[2], Cecilia Rumberia[2], Martin Norling[2,4], Richard P Bishop[2], Appolinaire Djikeng[3], Stephen J Kemp[2], Alan Orth[2], Robert A Skilton[3], Karl Ståhl[1,4] and Anne Fischer[2,5]

Abstract

Background: The rising demand for pork has resulted in a massive expansion of pig production in Uganda. This has resulted in increased contact between humans and pigs. Pigs can act as reservoirs for emerging infectious diseases. Therefore identification of potential zoonotic pathogens is important for public health surveillance. In this study, during a routine general surveillance for African swine fever, domestic pigs from Uganda were screened for the presence of RNA and DNA viruses using a high-throughput pyrosequencing method.

Findings: Serum samples from 16 domestic pigs were collected from five regions in Uganda and pooled accordingly. Genomic DNA and RNA were extracted and sequenced on the 454 GS-FLX platform. Among the sequences assigned to a taxon, 53% mapped to the domestic pig (*Sus scrofa*). African swine fever virus, Torque teno viruses (TTVs), and porcine endogenous retroviruses were identified. Interestingly, two pools (B and C) of RNA origin had sequences that showed 98% sequence identity to Ndumu virus (NDUV). None of the reads had identity to the class Insecta indicating that these sequences were unlikely to result from contamination with mosquito nucleic acids.

Conclusions: This is the first report of the domestic pig as a vertebrate host for Ndumu virus. NDUV had been previously isolated only from culicine mosquitoes. NDUV therefore represents a potential zoonotic pathogen, particularly given the increasing risk of human-livestock-mosquito contact.

Keywords: Metagenomics, Ndumu virus, Pigs, Reservoir, Zoonoses

Background

The rising demand for livestock products in Africa has resulted in an increased use of intensive pig production systems across the continent. Pigs are frequently preferred to other livestock species due to their relatively rapid growth rate and large litter sizes. In Uganda, the pig production industry is rapidly developing, as it has the potential to provide financial returns over a relatively short time [1]. However, intensification of pig production leads to increased contact between humans and domestic animals with possible public health consequences. Zoonoses such as Cysticercosis, Swine Influenza virus, Nipah virus, Menangle virus, porcine Hepatitis E, *Staphylococcus aureus* and *Streptococcus suis* have been reported in pigs and there is clear

potential for domestic swine to act as a reservoir for many emerging and re-emerging infectious diseases [2-7]. A complete picture of other potential zoonoses carried by domestic pigs that could potentially cross over to humans due to close confinement of pigs and humans, especially in smallholder pig farmers in Africa, is still lacking. Early identification of the pathogen spectrum and diagnosis of potential zoonoses in domestic animals in close contact with humans is therefore important for public health surveillance and development of early warning systems. Due to resource constraints, such surveillance is currently minimally implemented in developing countries, such as Uganda. Recent advances in next generation, high-throughput sequencing technologies offer rapid ways to identify and analyze the presence and diversity of emerging pathogens [8,9]. Viral metagenomics is therefore increasingly applied in veterinary epidemiology, as exemplified in the recent identification of a novel porcine boca-like

* Correspondence: cmasembe@zoology.mak.ac.ug
[1]Department of Biological Sciences, Makerere University, Kampala, Uganda
Full list of author information is available at the end of the article

virus [7] and Torque teno sus viruses 1 and 2 [10]. In this study, during a routine general surveillance for African swine fever, serum samples from 16 domestic pigs in five districts in Uganda (Figure 1) were analyzed using 454 pyrosequencing to characterize viral diversity.

Findings

Methodology

Serum samples from 16 domestic pigs were collected from five localities in Uganda (Figure 1). The sampling was carried out as part of a study on African swine fever (endemic in the region) from localities with suspected outbreaks of the disease. Serum samples from pigs originating from the same geographical locality were combined to generate a total of six pools from across the country (A-Moyo; B-Lira; C-Gulu; D-Mityana; E-Mpigi). The samples were filtered through a 0.22 μm filter followed by DNA and RNA extraction using a Qiamp DNA Mini kit (Qiagen) and TRIzol reagent (Invitrogen), respectively. DNA and RNA were amplified using the modified random priming mediated sequence independent single primer amplification (RP-SISPA) methodology [11]. Each amplified sample was further processed as described for shotgun library preparation in GS FLX 454 technology. The sequencing reads were trimmed to remove SISPA primers and barcodes, and only reads with a length greater than 50 bp were retained. Low complexity repeats were masked using Repeatmasker (*RepeatMasker Open-3.0*.1996-2010 http://www.repeatmasker.org) and sequences with more than 50% repeats were excluded. The sequences in each pool were assembled using the Newbler assembler version 2.5.3 with default settings (Roche. Genome Sequencer FLX Data Analysis Software Manual. *Mannheim, Germany: Roche Applied Science, 2007*). Contiguous sequences (contigs) and reads which did not assemble into contigs were categorized using BLASTN and BLASTX homology searches against the non-redundant nucleotide and amino acid databases from NCBI (version June 2011). Taxonomic classification of each contig/read was investigated using MEGAN 4.0 [12].

Results

A total of 289,038 reads with an average length of 175 nucleotides was obtained. After filtering for length and repeat content, 190,706 reads remained. Seventy-seven percent of the reads were assembled into contigs. Both BLASTN and BLASTX analyses gave similar results. Subsequently, the results that follow are from the BLASTN analysis.
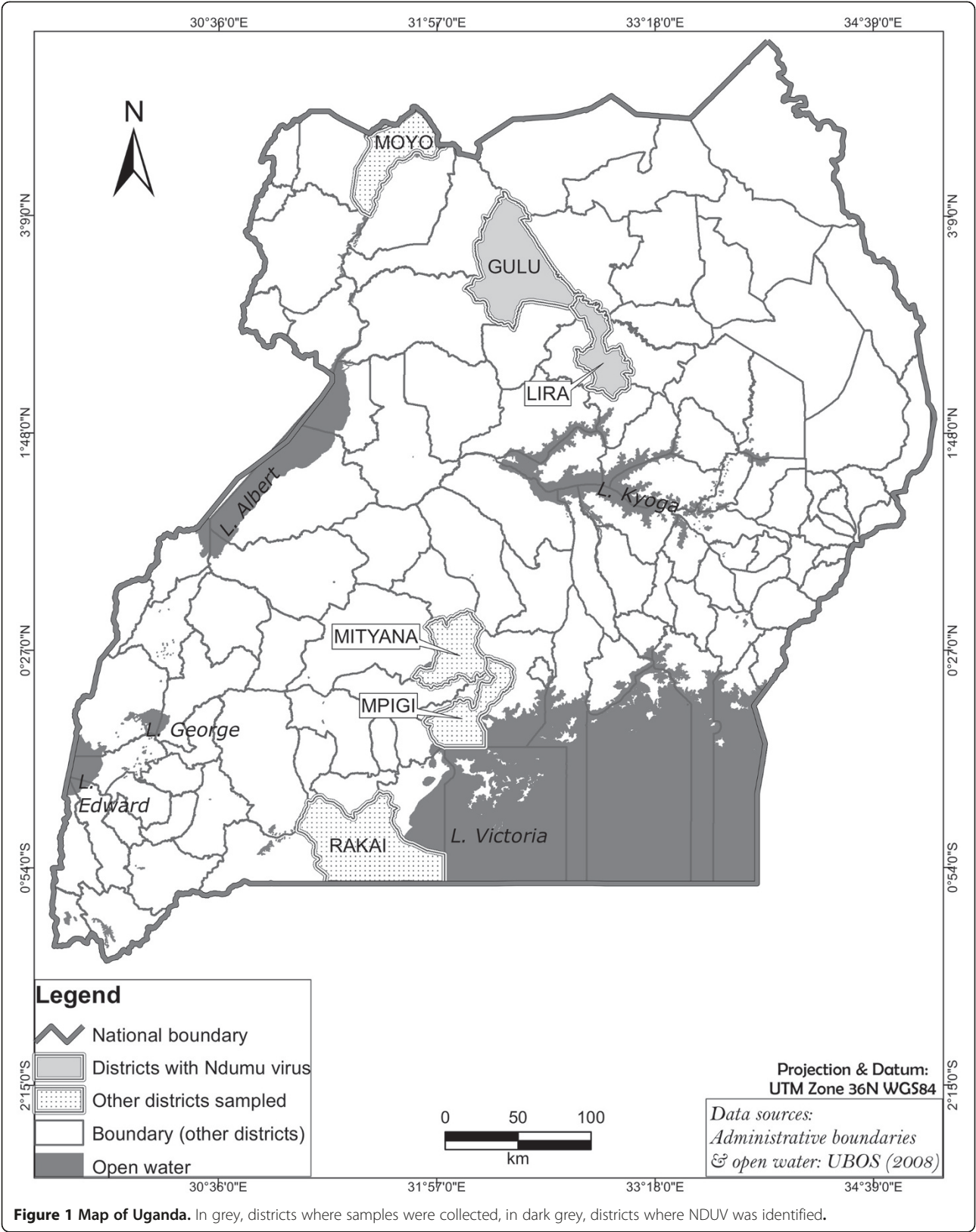
For 62% of all sequences, there was no significant match within GenBank as defined by the above criteria. Among the sequences assigned to a taxon, 52%, whether DNA or RNA, mapped to the domestic pig host (*Sus scrofa*). Thirty six percent of the sequences also exhibited similarity with other mammalian genomes (17% to the family *Bovidae* and 6.5% to human); these additional matches were most likely a consequence of the currently incomplete status of the porcine genome.

A small proportion (6.2%) of the sequences mapped to DNA and RNA viruses. For the DNA searches, besides African swine fever virus, which was identified in all pools, Torque teno viruses (TTVs) were identified in two pools (A and D). TTVs are ubiquitous species-specific viruses that are currently considered non-pathogenic and have been reported to infect swine with a high prevalence [10,13].

From the RNA pools analyses, three pig serum pools (A, C, & D) showed sequences with significant sequence similarity to diverse porcine endogenous retroviruses. Of particular interest are sequences that showed highly significant identity (98%) to NDUV and were present in pools B and C of pig RNA samples from Gulu and Lira districts, respectively. To validate this observation, the distribution of the sequence reads used earlier to build contigs from the two pools were determined by BLASTN analysis. RNA pool B had 6,228 sequence reads which comprised of 2% virus, 1% bacteria, 57% *Mammalia* and 40% unknown. RNA pool C had 64,583 reads which comprised of 5% virus, 1% bacteria, 51% *Mammalia* and 43% unknown (Additional file 1: Figure S1). The *Mammalia* genomes were from pig, human and *Bovidae*. Bacteria reads in both pools had identities of less than 30 bases at below 65% identity to *Eubacterium hallii* and *Xylella fastidiosa*. The bacteria sequences may therefore have occurred by chance and hence not significant explaining why they were lost on assembly of reads to contigs. No read showed a match to the class *Insecta*.

For the Ndumu virus (NDUV), a mapping assembly against the Ndumu genome sequence available in Genbank (NC_016959) using gsMapper was performed (Roche. Genome Sequencer FLX Data Analysis Software Manual. *Mannheim, Germany: Roche Applied Science, 2007*). Two contigs were built and the average coverage for each base was 10-fold. The mapped contigs were then masked: a minimum of two independent reads and a base quality of 20 was required for a base to be called. Masked sequences of the NDUV are available in GenBank under accession numbers JN989957-JN989958. The presence of NDUV was confirmed by amplification of 118 base pairs (bp) of the NDUV virus with NDUV-specific primers (Ndu-F1 GCCTACGTAGAACGTGCAGA and Ndu-R1 TGATGTTTCCCAACGTCACT). The PCR product was purified and sequenced on an ABI Prism 3700 DNA analyzer (Perkin-Elmer Applied Biosystems, Foster City, CA). The resulting sequences were

Figure 1 Map of Uganda. In grey, districts where samples were collected, in dark grey, districts where NDUV was identified.

compared to other members of the alphavirus genus from Genbank (AF069903, U73745, EF536323, AB032553, AY702913, GQ433358, X04129, AF369024, AF079456, NC_016962, NC_016959, M69205, HM147992). The protein coding regions (non-structural and structural proteins) of these sequences, and that of the NDUV sequences in this study were aligned using ClustalW [14]. In addition, a complete genome of 11688 bp of NDUV (GenBank NC_016959.1), and sequences of the Semliki Forest virus complex were added to the data set to construct phylogenetic trees. The alignment of the sequences is available as Additional file 2: Figure S2 and also as a complete file (Additional file 3: Sequence data). A neighbor-joining and maximum likelihood phylogenetic tree of the aligned protein sequences was constructed with MEGA 5.0 [15] using the Jones, Taylor and Thornton model of evolution and gamma distributed rates at sites incorporating 500 bootstrap replicates to assess the support of the phylogeny. Sequences of Venezuelan Equine Encephalitis Virus (VEEV) and Sindbis-Ocklebo virus were used as outgroups, based on previous phylogenetic studies [16]. The pig serum-derived RNA pool (B &C) sequences cluster with the currently available NDUV virus sequence as shown in Figure 2,

confirming that for the first time, this virus is detected in a vertebrate host, the domestic pig.

Discussion

In this study, a metagenomics approach was used to determine the variety of viruses in domestic pig (*Sus scrofa*) serum. In addition to the detection of mammalian sequences, it revealed the presences of some viruses for example Torque teno viruses (TTVs) and bacteria *Eubacterium hallii* and *Xylella fastidiosa* that have previously been found to occur regularly in pigs and plants, respectively. However, in the same study, a virus that has not been found before in pigs, the NDUMU (NDUV) virus, was detected. NDUV is a single stranded RNA arbovirus transmitted by mosquitoes and belonging to the *Togaviridae* family in the alphavirus genus. Very little is known about NDUV and its vertebrate hosts. It was isolated for the first time in South Africa in 1959 from *Mansonia uniformis* [17] and later in Kenya from *Aedes mcintoshi* and *A. ochraceus* [18]. Mice experimentally infected with NDUV do not survive the infection [17]. Although antibodies to the virus have been identified in humans from several African countries, no human morbidity or mortality has yet been
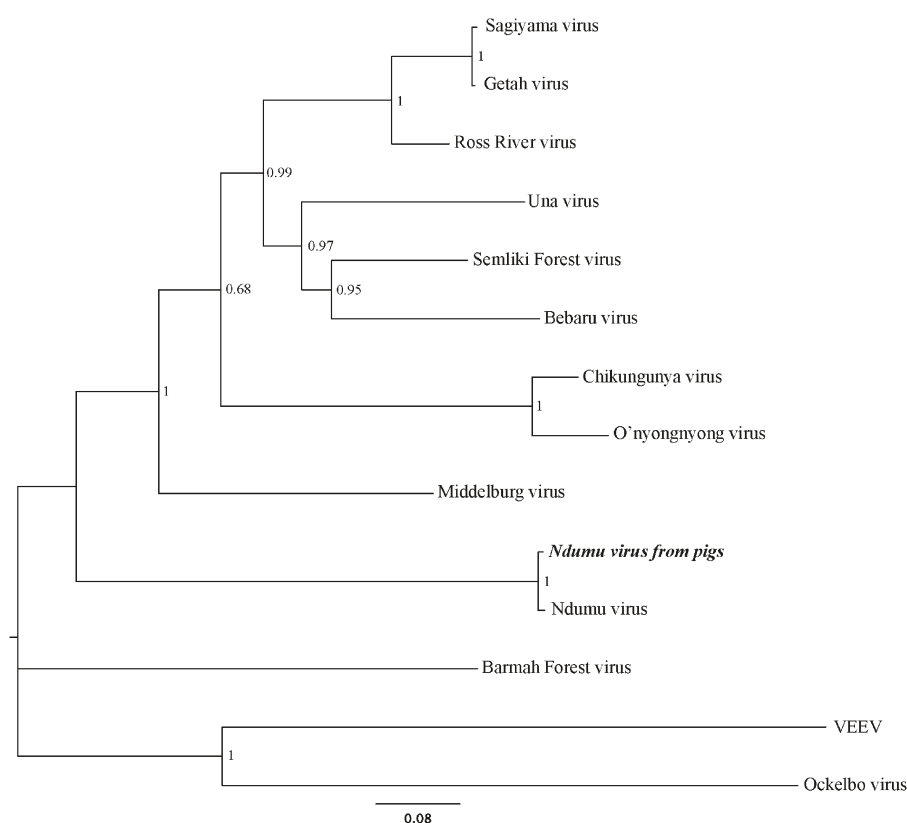

Figure 2 Maximum Likelihood phylogenetic tree of selected alphaviruses protein sequences. The NDUV from this study is in bold italics. Numbers on internal branches indicate bootstrap values for 500 replicates. VEEV: Venezuelan Equine Encephalitis Virus.

attributed to NDUV infection [17]. However the genus Alphavirus comprises at least 24 members [19], among which are many viruses, which cause diseases in humans and other animals. Chikungunya virus is one example of an Alphavirus that was responsible for recent severe outbreaks of human disease in Eastern Africa. In humans, the symptoms associated with Alphavirus infections range from fevers and rashes, to transient or debilitating arthritis, or encephalitis [20,21]. In this study, the domestic pig has for the first time been identified as a potential vertebrate host of NDUV. NDUV therefore represents a potential zoonotic agent, given the increasing risk of human-livestock-mosquito contact as the pig industry continues to intensify, and the pig population increases in Uganda. Our discovery indicates that a focused search for the virus using reverse-transcription PCR should now be performed in human communities associated with the domestic pig populations in which we have detected NDUV.

Additional files

Additional file 1: Figure S1. Taxonomic classification of sequence reads used to build contigs for RNA pools B (Gulu district) and C (Lira district) based on BLASTN (E-value <0.001) against genebank non-redundant database.

Additional file 2: Figure S2. Window showing alignment of amino acids of NDUV in this study with the selected alphaviruses used for phylogenetic reconstruction.

Additional file 3: Sequence data. Fasta format data file showing alignment of amino acids of NDUV in this study with the selected alphaviruses used for phylogenetic reconstruction.

Competing interests
The authors do hereby declare that they have no competing interests whatsoever in this scientific work.

Authors' contributions
CM conducted fieldwork, contributed to field study design, contributed to lab work, draft manuscript, and final manuscript preparation. GM, MO, CR performed the laboratory experiments and contributed to the data analysis and drafting of the manuscript. RPB, AD, SJK, RS contributed to study design and manuscript preparation. MN provided python scripts. AO installed and maintained the necessary programs on the server. KS coordinated the fieldwork and contributed to field study design, and final manuscript preparation. AF performed data analysis, wrote a manuscript draft and contributed to final manuscript preparation. All authors have read and approved the final manuscript.

Author details
[1]Department of Biological Sciences, Makerere University, Kampala, Uganda.
[2]International Livestock Research Institute (ILRI), Nairobi, Kenya. [3]Biosciences Eastern and Central Africa (BecA) –ILRI Hub Nairobi, Nairobi, Kenya. [4]Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden. [5]International Centre of Insect Physiology and Ecology, icipe, Nairobi, Kenya.

References
1. UBOS: *Uganda Bureau of Statistics.*: Statistical Abstract; 2011.
2. Banks M, Bendall R, Grierson S, Heath G, Mitchell J, Dalton H: **Human and porcine hepatitis E virus strains, United Kingdom.** *Emerg Infect Dis* 2004, **10**(5):953–955.
3. Myers KP, Olsen CW, Gray GC: **Cases of Swine influenza in humans: a review of the literature.** *Clin Infect Dis* 2007, **44**(8):1084–1088.
4. Parashar UD, Sunn LM, Ong F, Mounts AW, Arif MT, Ksiazek TG, *et al*: **Case-control study of risk factors for human infection with a new zoonotic paramyxovirus, Nipah virus, during a 1998-1999 outbreak of severe encephalitis in Malaysia.** *J Infect Dis* 2000, **181**(5):1755–1759.
5. Philbey AW, Kirkland PD, Ross AD, Davis RJ, Gleeson AB, Love RJ, *et al*: **An apparently new virus (family Paramyxoviridae) infectious for pigs, humans, and fruit bats.** *Emerg Infect Dis* 1998, **4**(2):269–271.
6. Smith TC, Harper AL, Nair R, Wardyn SE, Hanson BM, Ferguson DD, *et al*: **Emerging Swine zoonoses.** *Vector Borne Zoonotic Dis* 2011, **11**(9):1225–1234.
7. Blomstrom AL, Belak S, Fossum C, Mckillen J, Allan G, Wallgren P, Berg M: **Detection of a novel boca-like virus in the background of porcine circovirus type 2 induced postweaning multisytemic wasting syndrome.** *Virus Res* 2009, **146**:125–129.
8. Svraka S, Rosario K, Duizer E, van der Avoort H, Breitbart M, Koopmans M: **Metagenomic sequencing for virus identification in a public-health setting.** *J Gen Virol* 2010, **91**:2846–2856.
9. Mokili JL, Rohwer F, Dutilh BE: **Metagenomics and future perspectives in virus discovery.** *Curr Opin Virol* 2012, **2**:63–77.
10. Brink M, Ståhl K, Masembe C, Okurut AR, Berg M, Blomström AL: **First time molecular detection and phylogenetic relationships of Torque teno sus virus 1 and 2 in domestic pigs in Uganda: further evidence for a global distribution.** *Virol J* 2012, **9**:39. doi:10.1186/1743-422X-9-39.
11. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldblyum J, Sengamalay N, *et al*: **Viral genome sequencing by random priming methods.** *BMC Genomics* 2008, **9**:5.
12. Huson DH, Auch AF, Qi J, Schuster SC: **MEGAN analysis of metagenomic data.** *Genome Res* 2007, **17**(3):377–386.
13. McKeown NE, Fenaux M, Halbur PG, Meng XJ: **Molecular characterization of porcine TT virus, an orphan virus, in pigs from six different countries.** *Vet Microbiol* 2004, **104**(1–2):113–117.
14. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG, Clustal W, Clustal X: **version 2.0.** *Bioinformatics* 2007, **23**:2947–2948.
15. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: **MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol* 2011, **28**(10):2731–2739.
16. Forrester NL, Palacios G, Tesh RB, Savji N, Guzman H, Sherman M, Weaver SC, Lipkin WI: **Genome-scale phylogeny of the alphavirus genus suggests a marine origin.** *J Virol* 2012, **86**(5):2729–3.8.
17. Kokernot RH, McIntosh BM, Worth CB: **Ndumu virus, a hitherto unknown agent, isolated from culicine mosouitoes collected in northern Natal. Union of South Africa.** *AmJTrop Med Hyg* 1961, **10**:383–386.
18. Crabtree M, Sang R, Lutomiah J, Richardson J, Miller B: **Arbovirus surveillance of mosquitoes collected at sites of active Rift Valley fever virus transmission: Kenya, 2006-2007.** *J Med Entomol* 2009, **46**(4):961–964.

19. Weaver SC, Dalgarno L, Frey TK, Huang HV, Kinney RM, Rice CM, *et al*:
Family togaviridae. In *Virus taxonomy. Classification and nomenclature of
viruses. Seventh report of the International Committee on Taxonomy of Viruses.*
Edited by van Regenmortel CMF MHV, D H L, Bishop EBC, Estes MK, Lemon
SM, Maniloff J, Mayo MA, McGeogh CRP DJ, Wickner RB. San Diego:
Academic Press, Inc; 2000:879–889.
20. Tesh RB: **Arthritides caused by mosquito-borne viruses.** *Ann Rev Med*
1982, **33**:31–40.
21. Whitley RJ: **Viral encephalitis.** *New England J of Med* 1990, **323**:242–250.