

**Short
Communication****Metagenomic identification of a novel anellovirus in Pacific harbor seal (*Phoca vitulina richardsii*) lung samples and its detection in samples from multiple years**Terry Fei Fan Ng,¹ Elizabeth Wheeler,² Denise Greig,² Thomas B. Waltzek,^{3,4} Frances Gulland² and Mya Breitbart¹**Correspondence**

Mya Breitbart

mya@marine.usf.edu

¹University of South Florida College of Marine Science, 140 7th Avenue South, St Petersburg, FL 33701, USA²The Marine Mammal Center, Fort Cronkhite, Sausalito, CA 94965, USA³Department of Medicine and Epidemiology, School of Veterinary Medicine, University of California, Davis, CA 95616, USA⁴Department of Environmental and Global Health, College of Public Health and Health Professions, University of Florida, Gainesville, FL 32610, USA

To investigate viral pathogens potentially involved in a mortality event of 21 Pacific harbor seals (*Phoca vitulina richardsii*) in California in 2000, viral metagenomics was performed directly on lung samples from five individuals. Metagenomics revealed a novel seal anellovirus (SealAV), which clusters phylogenetically with anelloviruses from California sea lions and domestic cats. Using specific PCR, SealAV was identified in lung tissue from two of five animals involved in the 2000 mortality event, as well as one of 20 harbor seal samples examined post-mortem in 2008. The identification of SealAV in multiple years demonstrates that this virus is persistent in the harbor seal population. SealAV is the second anellovirus reported in the lungs of pinnipeds, suggesting that anellovirus infections may be common amongst marine mammals and that more research is needed to understand the roles of these viruses in marine mammal health and disease.

Received 14 December 2010

Accepted 7 March 2011

Emerging viral diseases can cause epidemics in wild animal populations, and surveillance for these pathogens is important for protecting both animal and human health (Kruse *et al.*, 2004; Kuiken *et al.*, 2005). However, diagnosing novel viral infections is difficult, because current viral identification techniques have limited capability for characterizing novel viruses. Diagnostic methods such as ELISA with antibodies for a specific virus, PCR or microarrays with primers designed for a specific virus or PCR with degenerate primers that amplify a closely related group of viruses are effective methods for detecting close relatives of known viruses (e.g. Symonds *et al.*, 2009; Wang *et al.*, 2002), but are limited for identification of new viruses. Viral particle purification and shotgun sequencing (viral metagenomics) has been developed as a method to discover novel viruses with limited sequence similarity to previously described viral families (reviewed by Delwart, 2007; Edwards & Rohwer, 2005). Recent advances in viral metagenomic techniques have enabled viral discovery directly from animal tissues, such as tumours (Ng *et al.*,

2009b) and lungs (Ng *et al.*, 2009a; Willner *et al.*, 2009), allowing the investigation of viruses involved in disease of humans or other animals.

Mass mortality of Pacific harbor seals (*Phoca vitulina richardsii*) occurs sporadically along the coast of California, but the causes of most of these events remain unknown (Nollens *et al.*, 2010). In 1997, approximately 90 harbor seals were found dead along a localized 10 mile (16 km) section of the central California coast at Point Reyes National Seashore, but no conclusive aetiological agent was identified. A similar event occurred in June 2000, when 21 adult harbor seals were found dead at the same location. Upon post-mortem examination, all animals had grossly abnormal lungs, and histological examination of three fresh cases revealed severe pneumonia. *Pseudomonas aeruginosa* was cultured from three of these seals (Gaffney *et al.*, 2008), and a coronavirus was characterized from one of the animals (Nollens *et al.*, 2010), but the identity of the overall causative agent of the mortality event is still uncertain.

This study used viral metagenomics to investigate DNA viruses in frozen lung samples from five of the seals

The GenBank/EMBL/DBJ accession number for the genome sequence of seal anellovirus TFFN/USA/2006 is HQ287751.

involved in the 2000 mortality event. Viral metagenomics was performed using previously published protocols (Breitbart & Rohwer, 2005; Ng *et al.*, 2009a, b). Briefly, lung tissue from the five seals was pooled and homogenized in sterile SM buffer (50 mM Tris/HCl, 10 mM MgSO₄, 0.1 M NaCl, pH 7.5). Host cells were removed through centrifugation at 10 000 *g* for 10 min, followed by filtration of the supernatant through a 0.22 μm filter. The filtrate was treated with 0.2 vols chloroform for 10 min and then incubated with 2.5 U DNase I μl⁻¹ for 3 h at 37 °C. DNA was extracted from the purified viral particles using the QIAamp MinElute Virus Spin kit (Qiagen) and amplified with the strand-displacement method of the GenomiPhi V2 DNA Amplification kit (GE Healthcare) according to the manufacturers' instructions. The GenomePlex Whole Genome Amplification kit (Sigma-Aldrich) was then used to fragment and further amplify the viral DNA, which was subsequently cloned into the pCR4 vector using TOPO TA cloning (Invitrogen). The resulting clones (*n*=174) were sequenced using Sanger sequencing and the metagenomic sequences were analysed using TBLASTX against the GenBank non-redundant database (Altschul *et al.*, 1990, 1997).

Analysis of the viral metagenome revealed sequences with amino acid identity to viruses in the family *Anelloviridae* (Carstens, 2010), which contain small, negative-sense, circular, single-stranded DNA genomes (Fauquet *et al.*, 2005). Anelloviruses are subgrouped into Torque teno virus (TTV), Torque teno mini virus (TTMV), Torque teno midi virus (TTMDV) and small anellovirus (SAV). Known hosts for anelloviruses include humans, non-human primates and domestic animals (Biagini *et al.*, 2007; Biagini, 2009; Hino & Miyata, 2007; Leary *et al.*, 1999). Recent metagenomic analyses have also identified anelloviruses in human blood (Breitbart & Rohwer, 2005; Jones *et al.*, 2005) and sea lions (Ng *et al.*, 2009a). To sequence the entire genome of the seal anellovirus (SealAV), strand-displacement amplification (TempliPhi; GE Healthcare) and inverse PCR was performed using primers SealAV-1585F (5'-GAACAGTCAGCAAAGAC-TCCA-3') and SealAV-1587R (5'-TTCTTCGGTAGACA-GAGTCTCC-3') designed from the metagenomic sequences. The PCR [containing 1 μM of each primer, 200 μM dNTPs, 1 U Red *Taq* DNA polymerase (Sigma-Aldrich), 1 × Red *Taq* reaction buffer and 5 μl target DNA in a 50 μl reaction] was amplified as follows: 95 °C for 5 min, 45 cycles of 94 °C for 1 min, 56 °C (temperature lowered by 0.2 °C per cycle) for 1 min and 72 °C for 1 min followed by 72 °C for 10 min. The resulting genome of SealAV was annotated using Geneious (Biomatters) and deposited into GenBank.

The circular genome of SealAV is 2164 nt long, with a G+C content of 47 mol%. The genome organization of SealAV is consistent with that of other anelloviruses, with three ORFs located on the negative-stranded genome (Fig. 1). ORF1 (1410 nt) does not display any significant nucleotide identity to any virus sequence using BLASTN, but

shares limited pairwise identity with the sea lion (*Zalophus californianus*) anellovirus ZcAV (57%), Torque teno felis virus Fc-TTV4 (54%) and feline anelloviruses PRA1 (54%) and PRA4 (55%) (Biagini *et al.*, 2007; Ng *et al.*, 2009a; Okamoto *et al.*, 2002) by optimal pairwise alignment of the complete ORF1 nucleotide sequence in BioEdit (Hall, 1999). ORF1 encodes 469 amino acids and has weak amino acid sequence identity to the sea lion anellovirus ZcAV (21%), Torque teno felis virus Fc-TTV4 (18%) and feline anelloviruses PRA1 (21%) and PRA4 (18%) based on optimal pairwise amino acid alignment in BioEdit. The N terminus of ORF1 contains an arginine-rich region, with 26 arginine residues in the first 63 amino acids (41%), which is common in single-stranded DNA animal viruses, including anelloviruses, circoviruses and gyroviruses (Ng *et al.*, 2009a). A nuclear localization signal was predicted for ORF1 using PREDICTNLS (Cokol *et al.*, 2000) based on the amino acid sequence KRVRFR RGDIRRKRL near the 3' end. ORF2 encodes 111 amino acids and shares no amino acid sequence identity with proteins in GenBank. ORF2 lacks the motif W-X₇-H-X₃-C-X₁-C-X₅-H that is conserved among TTVs, TTMVs and TTMDVs (Hijikata *et al.*, 1999; Ninomiya *et al.*, 2007, 2009; Okamoto *et al.*, 2000; Takahashi *et al.*, 2000). ORF3 encodes 204 amino acids, and shares weak amino acid sequence identity (16%) with ZcAV ORF3 only, indicating that SealAV could be more closely related to ZcAV than to

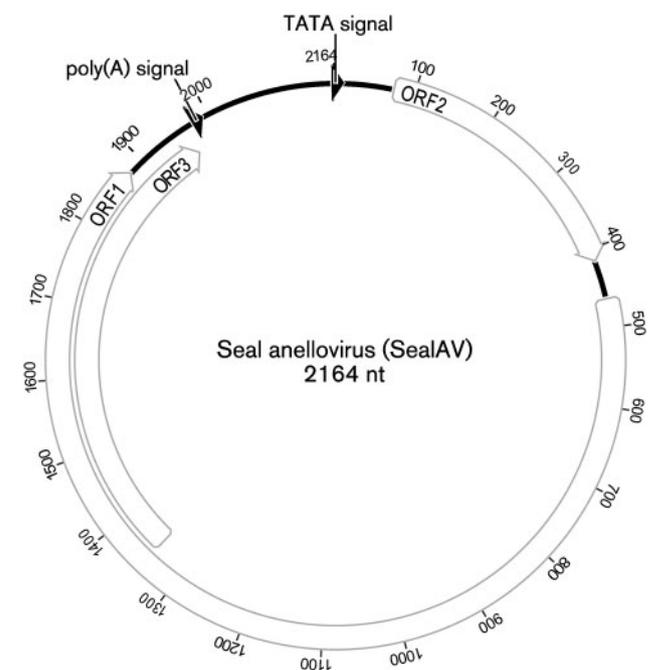


Fig. 1. Genomic organization of the seal anellovirus (SealAV). ORF1 shares weak amino acid sequence identity with ORF1 of the California sea lion anellovirus ZcAV and feline anelloviruses Fc-TTV4, PRA1 and PRA4. ORF2 has no amino acid sequence identity to proteins in GenBank. ORF3 shares weak amino acid identity with ZcAV.

other anelloviruses. The SealAV and ZcAV genomes share only a low level of genome-wide nucleotide identity (59%), making it very unlikely that laboratory contamination contributed to the relatedness of these viruses. The SealAV genome shares extremely low nucleotide sequence identity with all known anelloviruses, preventing its detection with degenerate PCR (such as Okamoto *et al.*, 2002) or microarrays designed based on known genomes.

To determine the relationship of SealAV to other anelloviruses, a phylogram was created based on the amino acid sequence encoded by ORF1 (Biagini *et al.*, 2007; Ng *et al.*, 2009a). Deduced amino acid sequences were aligned (1161 amino acid characters including gaps for 44 taxa) using Mafft 5.8 (Katoh *et al.*, 2005) with the E-INS-I alignment strategy and the following parameters: scoring matrix BLOSUM62, gap open penalty 1.53 and offset value 0. Bayesian inference trees were constructed using MrBayes version 3.1.2 (Huelsenbeck & Ronquist, 2001) using the top scoring model (WAG+I+G+F) recommended by ProtTest 1.4 (Abascal *et al.*, 2005) under the Bayesian information criterion. The Markov chain was run for a maximum of 10 million generations, with a stopping rule implemented so that the analysis would halt when the

average deviation of the split frequencies was less than 0.001%. Four independent analyses were conducted, each with one cold and three heated chains with the default heating parameter (temperature=0.2). Every 50 generations were sampled and the first 25% of MCMC samples were discarded as burn-in. In addition to the Bayesian analysis, a neighbour-joining phylogenetic tree was created based on the entire nucleotide sequence of ORF1 (Biagini *et al.*, 2007) and the topology of the resulting tree (not shown) was nearly identical to that of the Bayesian inference tree.

The anellovirus tree topology was consistent with results of earlier studies (Biagini, 2009; Jones *et al.*, 2005) and revealed that SealAV is a novel anellovirus species that belongs to a clade that contains the sea lion anellovirus and three feline anelloviruses (Fig. 2). This clade is part of a larger cluster of viruses from domestic and wild animals distinct from the TTV, TTMV, TTMDV and SAV clades that primarily infect humans (Biagini, 2009). Although phylogenetic analysis of ORF1 did not resolve the relationships of the two pinniped anelloviruses to each other or to the feline anelloviruses, the top TBLASTX hit for SealAV ORFs 1 and 3 was ZcAV, suggesting relatedness between

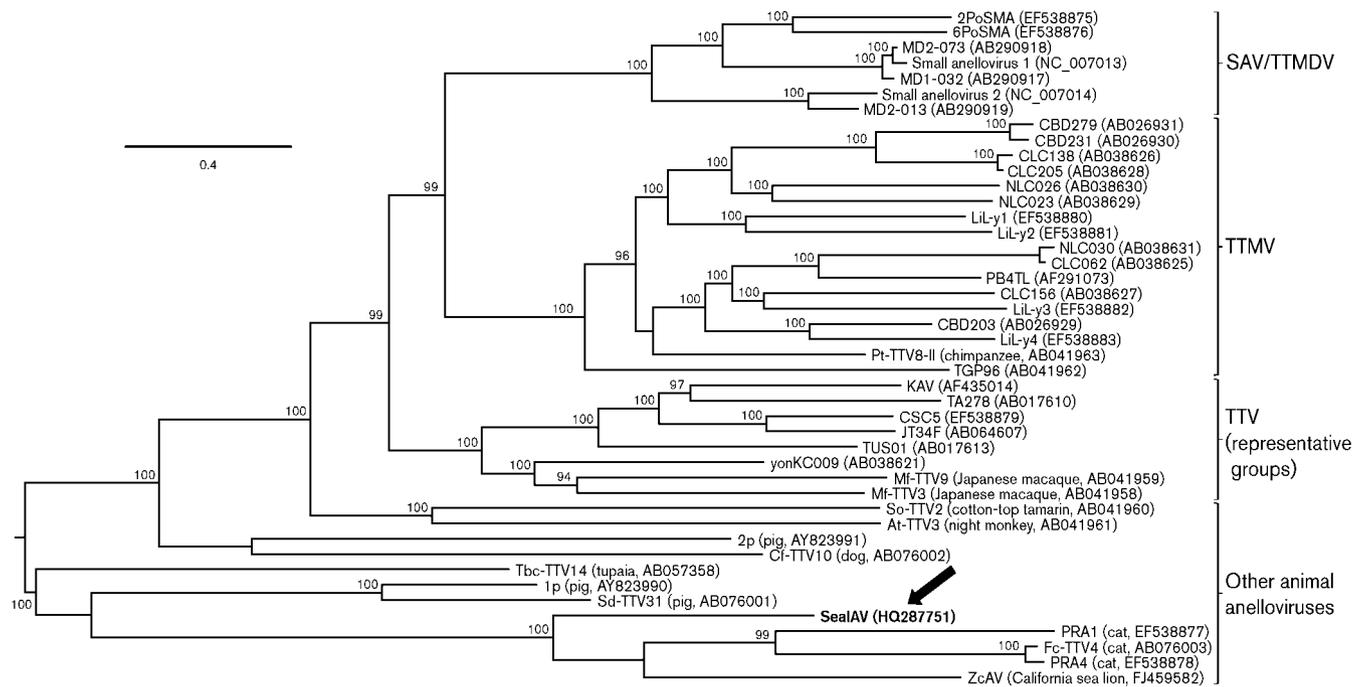


Fig. 2. Phylogram based on the full-length amino acid sequence of ORF1 depicting relationships among the members of the family *Anelloviridae*. The newly discovered harbor seal anellovirus (SealAV) is indicated by an arrow and joins a clade of anelloviruses that infect domestic cats and California sea lions. This clade is distinct from the TTV, TTMDV, SAV and TTMV clades that primarily infect humans and primates. The host is indicated in parentheses for any non-human sequences and GenBank accession numbers are shown. Numbers above each node represent posterior probabilities (values >90% shown) of the Bayesian analysis. Branch lengths are based on the number of inferred amino acid substitutions, as indicated by the bar (0.4 substitutions per amino acid position). The tree was created in FigTree 1.3.1 (Rambaut, 2010) using the midpoint rooting feature.

the two pinniped anelloviruses. In contrast to the clustering observed for the pinniped and feline anelloviruses, the host seals (phocids) and sea lions (otariids) are more closely related to dogs (canids) than to cats (felids) (Eizirik *et al.*, 2010). Therefore, co-diversification between these anelloviruses and their hosts is not supported.

Three distinct species of mRNA have been observed for human anelloviruses and most of the animal anelloviruses through transcriptional studies (Kamahora *et al.*, 2000; Okamoto *et al.*, 2001), but only the second splicing event was detected in the genomes of the sea lion anellovirus and the tupaia TTV (Ng *et al.*, 2009a; Okamoto *et al.*, 2001). Analysis of potential donor and acceptor sites for the SealAV genome using GeneSplicer (<http://cbcb.umd.edu/software/GeneSplicer/>) only recognized the second and third splicing events; future work needs to examine transcriptional splicing of SealAV experimentally (de Villiers *et al.*, 2009).

To investigate the prevalence of SealAV, PCR was performed using DNA from individual lung samples. Total DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen). Specific PCR to detect SealAV was executed using the primers SealAV-1068F (5'-TCTCC-ACACAGGACAACAGC-3') and SealAV-1313R (5'-GAG-ACCAGACCACCATGGTC-3'), which have a product size of 246 bp. The PCR was the same as described above. Two of five seals from the 2000 mortality event tested positive for SealAV, and positive PCR products were sequenced for confirmation. One of the anellovirus-positive seals was previously shown to contain the harbor seal coronavirus (Nollens *et al.*, 2010), demonstrating co-infection of this individual seal with both viruses.

In 2008, 13 harbor seal pups and three yearlings that stranded along the central California coast were admitted to the Marine Mammal Center for rehabilitation. Despite treatment, these animals died, and representative lung samples from these animals were examined histologically and archived at -70°C for virology studies. In addition, four seal carcasses discovered along the coast in the same period were collected for analysis. PCR for SealAV was performed on these 20 lung samples as described above, revealing one positive result in a post-mortem harbor seal pup without pneumonia. SealAV could not be detected by PCR in the blood of this anellovirus-positive pup. The difference in prevalence between the samples from 2000 and 2008 could be due to the age of harbor seals tested (adults vs pups), temporal variation, health status of the seals or many other host or environmental conditions. Nevertheless, these results confirm that SealAV was present in harbor seals sampled 8 years apart, suggesting that this virus is persistent in the seal population. It should also be noted that PCR data might underestimate the actual number of anellovirus infections in wild seals, because the immune systems of some animals could have cleared the viral infection at the time of sample collection, resulting in a false-negative PCR result.

It has been suggested that TTV can be involved in respiratory disease complexes or enhance the effects of other pathogens in pigs and humans (Davidson & Shulman, 2008; Ellis *et al.*, 2008; Krakowka & Ellis, 2008; Maggi *et al.*, 2003, 2004; Taira *et al.*, 2009), but the definitive pathogenicity of anelloviruses remains unknown. Since SealAV was only recovered in a small proportion of the seals from the mortality event and no consistent findings were noted in necropsies of the positive seals, SealAV is probably not a causative agent of the mortality event. However, as the second report of anelloviruses in the lungs of pinnipeds, this work suggests that anellovirus infections may be common amongst marine mammals and that more research is needed to understand the roles of these viruses in marine mammal health and disease. Unfortunately, the fact that these viruses cannot be detected in the blood of infected animals by PCR presents a technical limitation to determining viral prevalence and pathogenicity, since lung biopsies cannot be obtained from healthy individuals. Future studies aimed at developing an antibody-based serological test for anelloviruses will be useful for determining the exposure of wild pinniped populations to these viruses.

In conclusion, viral metagenomics enabled the discovery of a novel seal anellovirus from a mortality event of harbor seals in California in 2000, and the same virus was detected in a dead seal in 2008. Future studies need to determine the diversity of anelloviruses that infect marine mammals and examine the roles of anelloviruses in the health of these animals.

Acknowledgements

This project was funded by a grant to M. B. from the Alfred P. Sloan Foundation (BR-4772). T.F.F.N. was funded by the Gulf Oceanographic Charitable Trust Fellowship and the William and Elsie Knight Oceanographic Fellowship. F.G., E.W. and D.G. were supported with funds from the NMFS John H. Prescott Marine Mammal Stranding Assistance Program. Samples were collected under MMPA permit no. 932-19051MA-009526. Thanks to Darren Dunlap for assistance in sample processing.

References

- Abascal, F., Zardoya, R. & Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104–2105.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Biagini, P. (2009). Classification of TTV and related viruses (anelloviruses). *Curr Top Microbiol Immunol* **331**, 21–33.
- Biagini, P., Uch, R., Belhouche, M., Attoui, H., Cantaloube, J. F., Brisbarre, N. & de Micco, P. (2007). Circular genomes related to anelloviruses identified in human and animal samples by using a combined rolling-circle amplification/sequence-independent single primer amplification approach. *J Gen Virol* **88**, 2696–2701.

- Breitbart, M. & Rohwer, F. (2005). Method for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. *Biotechniques* **39**, 729–736.
- Carstens, E. B. (2010). Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2009). *Arch Virol* **155**, 133–146.
- Cokol, M., Nair, R. & Rost, B. (2000). Finding nuclear localization signals. *EMBO Rep* **1**, 411–415.
- Davidson, I. & Shulman, L. M. (2008). Unraveling the puzzle of human anellovirus infections by comparison with avian infections with the chicken anemia virus. *Virus Res* **137**, 1–15.
- Delwart, E. L. (2007). Viral metagenomics. *Rev Med Virol* **17**, 115–131.
- de Villiers, E.-M., Kimmel, R., Leppik, L. & Gunst, K. (2009). Intragenomic rearrangement in TT viruses: a possible role in the pathogenesis of disease. *Curr Top Microbiol Immunol* **331**, 91–107.
- Edwards, R. A. & Rohwer, F. (2005). Viral metagenomics. *Nat Rev Microbiol* **3**, 504–510.
- Eizirik, E., Murphy, W. J., Koepfli, K. P., Johnson, W. E., Dragoo, J. W., Wayne, R. K. & O'Brien, S. J. (2010). Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Mol Phylogenet Evol* **56**, 49–63.
- Ellis, J. A., Allan, G. & Krakowka, S. (2008). Effect of coinfection with genogroup 1 porcine torque teno virus on porcine circovirus type 2-associated postweaning multisystemic wasting syndrome in gnotobiotic pigs. *Am J Vet Res* **69**, 1608–1614.
- Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. & Ball, L. A. (editors) (2005). *Virus Taxonomy: Classification and Nomenclature of Viruses. Eighth Report of the International Committee on Taxonomy of Viruses*. San Diego: Elsevier Academic Press.
- Gaffney, P., Colegrove, K., Gulland, F., Byrne, B., Jang, S. & Edgar, K. (2008). Pathologic, microbiologic and epidemiologic characterization of *Pseudomonas* sp. in California sea lions (*Zalophus californianus*) and Pacific harbor seals (*Phoca vitulina*). In *Proceedings of the 39th Annual Conference of the International Association for Aquatic Animal Medicine*, 10–14 May 2008, Pomezia, Italy, p. 100. Edited by T. Reidarson. Rome: IAAAM.
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**, 95–98.
- Hijikata, M., Takahashi, K. & Mishiro, S. (1999). Complete circular DNA genome of a TT virus variant (isolate name SANBAN) and 44 partial ORF2 sequences implicating a great degree of diversity beyond genotypes. *Virology* **260**, 17–22.
- Hino, S. & Miyata, H. (2007). Torque teno virus (TTV): current status. *Rev Med Virol* **17**, 45–57.
- Huelsenbeck, J. P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755.
- Jones, M. S., Kapoor, A., Lukashov, V. V., Simmonds, P., Hecht, F. & Delwart, E. (2005). New DNA viruses identified in patients with acute viral infection syndrome. *J Virol* **79**, 8230–8236.
- Kamahora, T., Hino, S. & Miyata, H. (2000). Three spliced mRNAs of TT virus transcribed from a plasmid containing the entire genome in COS1 cells. *J Virol* **74**, 9980–9986.
- Katoh, K., Kuma, K., Toh, H. & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**, 511–518.
- Krakowka, S. & Ellis, J. A. (2008). Evaluation of the effects of porcine genogroup 1 torque teno virus in gnotobiotic swine. *Am J Vet Res* **69**, 1623–1629.
- Kruse, H., Kirkemo, A.-M. & Handeland, K. (2004). Wildlife as source of zoonotic infections. *Emerg Infect Dis* **10**, 2067–2072.
- Kuiken, T., Leighton, F. A., Fouchier, R. A. M., LeDuc, J. W., Peiris, J. S. M., Schudel, A., Stöhr, K. & Osterhaus, A. D. M. E. (2005). Public health. Pathogen surveillance in animals. *Science* **309**, 1680–1681.
- Leary, T. P., Erker, J. C., Chalmers, M. L., Desai, S. M. & Mushahwar, I. K. (1999). Improved detection systems for TT virus reveal high prevalence in humans, non-human primates and farm animals. *J Gen Virol* **80**, 2115–2120.
- Maggi, F., Pifferi, M., Fornai, C., Andreoli, E., Tempestini, E., Vatteroni, M., Presciuttini, S., Marchi, S., Pietrobelli, A. & other authors (2003). TT virus in the nasal secretions of children with acute respiratory diseases: relations to viremia and disease severity. *J Virol* **77**, 2418–2425.
- Maggi, F., Pifferi, M., Tempestini, E., Lanini, L., De Marco, E., Fornai, C., Andreoli, E., Presciuttini, S., Vatteroni, M. L. & other authors (2004). Correlation between Torque tenovirus infection and serum levels of eosinophil cationic protein in children hospitalized for acute respiratory diseases. *J Infect Dis* **190**, 971–974.
- Ng, T. F. F., Suedmeyer, W. K., Wheeler, E., Gulland, F. & Breitbart, M. (2009a). Novel anellovirus discovered from a mortality event of captive California sea lions. *J Gen Virol* **90**, 1256–1261.
- Ng, T. F. F., Manire, C., Borrowman, K., Langer, T., Ehrhart, L. & Breitbart, M. (2009b). Discovery of a novel single-stranded DNA virus from a sea turtle fibropapilloma by using viral metagenomics. *J Virol* **83**, 2500–2509.
- Ninomiya, M., Takahashi, M., Shimosegawa, T. & Okamoto, H. (2007). Analysis of the entire genomes of fifteen torque teno midi virus variants classifiable into a third group of genus Anellovirus. *Arch Virol* **152**, 1961–1975.
- Ninomiya, M., Takahashi, M., Hoshino, Y., Ichiyama, K., Simmonds, P. & Okamoto, H. (2009). Analysis of the entire genomes of torque teno midi virus variants in chimpanzees: infrequent cross-species infection between humans and chimpanzees. *J Gen Virol* **90**, 347–358.
- Nollens, H. H., Wellehan, J. F., Archer, L., Lowenstine, L. J. & Gulland, F. M. D. (2010). Detection of a respiratory coronavirus from tissues archived during a pneumonia epizootic in free-ranging Pacific harbor seals *Phoca vitulina richardsii*. *Dis Aquat Organ* **90**, 113–120.
- Okamoto, H., Nishizawa, T., Tawara, A., Peng, Y., Takahashi, M., Kishimoto, J., Tanaka, T., Miyakawa, Y. & Mayumi, M. (2000). Species-specific TT viruses in humans and nonhuman primates and their phylogenetic relatedness. *Virology* **277**, 368–378.
- Okamoto, H., Nishizawa, T., Takahashi, M., Tawara, A., Peng, Y. H., Kishimoto, J. & Wang, Y. (2001). Genomic and evolutionary characterization of TT virus (TTV) in tupaia and comparison with species-specific TTVs in humans and non-human primates. *J Gen Virol* **82**, 2041–2050.
- Okamoto, H., Takahashi, M., Nishizawa, T., Tawara, A., Fukai, K., Muramatsu, U., Naito, Y. & Yoshikawa, A. (2002). Genomic characterization of TT viruses (TTVs) in pigs, cats and dogs and their relatedness with species-specific TTVs in primates and tupaia. *J Gen Virol* **83**, 1291–1297.
- Rambaut, A. (2010). FigTree, version 1.3.1. Institute of Evolutionary Biology, University of Edinburgh.
- Symonds, E. M., Griffin, D. W. & Breitbart, M. (2009). Eukaryotic viruses in wastewater samples from the United States. *Appl Environ Microbiol* **75**, 1402–1409.
- Taira, O., Ogawa, H., Nagao, A., Tuchiya, K., Nunoya, T. & Ueda, S. (2009). Prevalence of swine Torque teno virus genogroups 1 and 2 in Japanese swine with suspected post-weaning multisystemic wasting syndrome and porcine respiratory disease complex. *Vet Microbiol* **139**, 347–350.

Takahashi, K., Hijikata, M., Samokhvalov, E. I. & Mishiro, S. (2000). Full or near full length nucleotide sequences of TT virus variants (Types SANBAN and YONBAN) and the TT virus-like mini virus. *Intervirology* **43**, 119–123.

Wang, D., Coscoy, L., Zylberberg, M., Avila, P. C., Boushey, H. A., Ganem, D. & DeRisi, J. L. (2002). Microarray-based detection and

genotyping of viral pathogens. *Proc Natl Acad Sci U S A* **99**, 15687–15692.

Willner, D., Furlan, M., Haynes, M., Schmieder, R., Angly, F. E., Silva, J., Tammadoni, S., Nosrat, B., Conrad, D. & Rohwer, F. (2009). Metagenomic analysis of respiratory tract DNA viral communities in cystic fibrosis and non-cystic fibrosis individuals. *PLoS ONE* **4**, e7370.