LUND UNIVERSITY

Characterization of the complete genomes of Camelus dromedarius papillomavirus types 1 and 2

Ure, Agustin; Elfadl, A. K.; Khalafalla, A. I.; Gameel, A. A. R.; Dillner, Joakim; Forslund, Ola

Link to publication

Characterization of complete genomes of Camelus dromedarius papillomavirus type 1 and 2

Ure, A. E. *, Elfadl, A.K. †, Khalafalla, A. I. ‡, Gameel, A. A. R. †, Dillner, J. * and O. Forslund *

* Department of Laboratory Medicine, Medical Microbiology, Skåne University Hospital, Lund University, Malmö, Sweden.

† Department of Pathology, Faculty of Veterinary Medicine, University of Khartoum, P. O. box 32, Khartoum North, Sudan

‡ Department of Microbiology, Faculty of Veterinary Medicine, University of Khartoum, P. O. box 32, Khartoum North, Sudan.

Corresponding author: Ola Forslund, ola.forslund@med.lu.se

Running title: Novel papillomaviruses isolated from dromedaries.

Keywords: Papillomavirus, Papillomaviridae, Arabic camel, dromedary, complete genome, CdPV1, CdPV2, one-humped.

Number or words in main text: 3934

Number of words in Summary: 169

Number of tables: 2

Number of figures: 3

[1]

Summary

Camel papillomatosis has been previously described but the genome of the suspected papillomavirus (PV) has not been identified. An outbreak of papillomatosis occurred in a dromedary farm of 55 animals in Sudan during August 2009. The disease was only present in young animals about 3-7 months old, of which 44 % (11/25) were affected with lesions mainly on the lips and lower jaw. Here we report for the first time the complete genomes of *Camelus dromedarius* papillomavirus type 1 (CdPV1) and 2 (CdPV2), isolated from a cauliflower-like nodule and a round oval raised nodule, respectively. Pairwise comparisons of their L1 nucleotide sequences revealed 69.2 % identity, and phylogenetic analyses suggest that these two PV-types are grouped within the genus *Deltapapillomavirus*. Both viruses were isolated from fibropapillomas, although no putative E5 proteins homologous to that of bovine papillomavirus type 1 were identified. The genetic information will be useful for evolutionary studies of the family *Papillomaviridae* as well as for the development of diagnostic methods for surveillance of the disease in dromedaries.

Introduction

The family *Papillomaviridae* is composed of double stranded circular DNA viruses that are widespread among vertebrates. These viruses are known to infect the epithelial tissues of reptiles, birds, and mammals, and almost all papillomaviruses are strictly specific to their natural host and do not infect even closely related species (Bernard *et al.*, 2010; Chow *et al.*, 2010).

Dromedaries (*Camelus dromedarius*), also known as Arabic or one-humped camels, are the most abundant camelids in the world, and a very important source of meat and milk in arid regions, where they also serve as working animals (Khalafalla *et al.*, 1998). In Sudan, the current dromedary population is estimated to be 4.6 million (Sudan Ministry of Animal Resources & Fisheries, 2010) and is located in the arid and semi-arid parts of the country, where the annual rainfall is less than 350 mm. The majority of the described camel papillomatosis cases occurred in the late rainy season (July-September) coinciding with outbreaks of camel contagious ecthyma and camel pox (Khalafalla, 1998; Munz *et al.*, 1990). Papillomatosis has also been shown in South American camelids (llama and alpaca) (Schulman *et al.*, 2003). The first report of papillomatosis in dromedaries was published in 1990 (Munz *et al.*, 1990), but the genetic characteristics of the suspected papillomavirus (PV) have not yet been described.

The aim of the present study is to isolate and characterize the complete genome of the dromedary papillomavirus.

Results

Papillomatosis in a herd of dromedaries

An outbreak of skin tumours occurred in a camel farm of 55 animals in Al Qutaynah locality, about 83 km south of Khartoum during August 2009. The disease occurred in

young animals about 3-7 months old, of which 44 % (11/25) were affected with warty lesions. The lesions were mainly located on the lips and lower jaw (see table 3 of supplementary material). Two types of skin lesions were identified. Only one animal demonstrated a cauliflower-like nodule (Fig. 1a), whereas the majority of animals exhibited fissured round or oval raised nodules (Fig. 1b). The lesions were dark grey, about 0.5-1 cm in length and 1-2 cm in width. Histopathological sections showed multiple papillary proliferations covered with keratinized epithelium and downgrowth of rete ridges for both types of lesions (Fig. 1c and d). There was acanthosis with karyopyknosis and cytoplasmic vacuolations in the stratum spinosum cells and hyperkeratosis in the upper layers. In addition, subepithelial fibrosis was also observed in sections from both dromedaries. No inclusions could be detected in squamous cells. Accordingly, the lesions were diagnosed as fibropapillomas.

Characterization of complete papillomavirus genomes

DNA was extracted from a cauliflower-like nodule from the lip of a male calf (Fig. 1a) and from a round oval raised nodule from the lip of another male calf (Fig. 1b). Both animals were less than eight months old. Amplicons were detected by PCR with FA-primers (FAP 59 and FAP64) (Forslund *et al.*, 1999). Sequencing of the product (ca. 480 bp) and subsequent analysis with BLASTn revealed subgenomic fragments of two novel putative papillomaviruses related to bovine and cervid PVs. Specific primers were then designed for each PV in a back-to-back orientation. After amplification by PCR, both complete genomes were cloned.

Each of the samples contained a unique complete PV sequence (Fig. 2) and comparisons of their L1 ORF revealed they had a 69.2 % nucleotide identity to each other. The closest related papillomavirus of the first isolate was BPV1 (66.3 %), while the second

closest were BPV1 and BPV2 (both 64.6 %), justifying the naming of the genomes as *Camelus dromedarius* papillomavirus type 1 and 2 (CdPV1 and CdPV2), correspondingly (Table 1). The complete genomes of CdPV1 and CdPV2 contain 7679 bp and 7906 bp with a GC content of 45.8 % and 47.2 %, respectively.

In both genomes, seven ORFs were identified (Fig. 2). The putative E6 protein from both dromedary papillomaviruses manifested two zinc-finger domains (Barbosa *et al.*, 1989), which are conserved among papillomaviruses and are present in all artiodactyl (even-toed ungulates) papillomaviruses currently known (Alberti *et al.*, 2010; Erdelyi *et al.*, 2008; Tomita *et al.*, 2007). Similarly, there was one zinc-finger domain in the E7 proteins but no retinoblastoma tumour suppressor (pRB) binding domain (L-X-C-X-E) (Dahiya *et al.*, 2000). In the E1 protein, the superfamily 3 ATP-dependent helicase domain (Iyer *et al.*, 2004; Liu *et al.*, 2010) was found in both viruses. The E2 proteins had the typical C-terminal DNA-binding domain and the N-terminal trans-activation domain (Hegde, 2002; Hegde *et al.*, 1992).

For both CdPV1 and CdPV2, the E4 ORF did not contain a start codon. The derived E4 protein is known to be generated by RNA splicing, generating an E1^E4 fusion product with approximately five amino acids from the E1 N-terminal end (Doorbar *et al.*, 1986). Based on BPV1 sequence information about donor and acceptor splicing sites (Doorbar & Myers, 1996; Yang *et al.*, 1985), corresponding sites were identified in E4 for both dromedary papillomaviruses.

No obvious E5 ORF could be identified in the region between E2 and L2, although a short ORF was present in both CdPVs (Fig. 2). Additional analysis of the conceptual translation of the short ORFs revealed in both cases a hydrophobic protein of 54 and 48 aa, with grand average values of hydropathy (GRAVY) of 1.498 and 1.077 for CdPV1 and CdPV2 ORFs, respectively. In addition, transmembrane regions were predicted to

be present in the putative protein from the short ORF of both CdPV1 and CdPV2 (data not shown).

During the sequencing of CdPV1 it was observed that the L2 ORF contained a region of eleven consecutive adenines (nt 5502). This was confirmed by additional sequencing of two different clones as well as amplicons from the original DNA sample (in both orientations). Nevertheless, it must be noted that the original clone contained only ten adenines, which caused a frameshift and a consequent overlap between the L2 and L1 ORF.

In the putative L1 protein, a nuclear localization signal ubiquitous for papillomaviruses (Zhou *et al.*, 1991) was found in both CdPVs (CdPV1: aa 481-497, CdPV2: 480-495).

Immediately downstream of the L1 ORF, a polyadenylation site (AATAAA) was identified in the CdPV1 (nt 7085) genome whereas a double polyadenylation site (AATAAATAAA) was found in the CdPV2 genome (nt 7145).

The upstream regulatory region (URR), also known as the long control region, was 622 bp for CdPV1 and 789 bp for CdPV2. The URR of CdPV1 demonstrated six E2 binding sites (E2BS) [ACC(N$_6$)GGT] (Androphy *et al.*, 1987; Li *et al.*, 1989). In the URR of CdPV2, seven E2BS were found as well as one in the 3' region of L1. Also, two putative E1 binding sites (AACAAT) (Chen & Stenlund, 2001) were found in the URR of CdPV1, and three in that of CdPV2. The origin of replication (*ori*) was identified for both CdPVs, homologous to the *ori* from BPV1, in which an AT-rich region is in close proximity to a dyad of E1 binding sites and an E2 binding site (Ustav *et al.*, 1993; Ustav & Stenlund, 1991; Yang & Botchan, 1990; Yang *et al.*, 1991). In both cases, the double E1 binding site of the *ori* was centred in an *HpaI* site (CdPV1: nt 7511, CdPV2: nt 7732), similarly to BPV1 (Spalholz *et al.*, 1993).

The phylogenetic relationships of CdPV1 and CdPV2 were established based on multiple alignments of their L1 ORFs, along with those of 71 prototype species of PVs. The phylogenetic analysis demonstrated that CdPV1 and CdPV2 were positioned on a separate branch between genera *Deltapapillomavirus* and *Epsilonpapillomavirus* (Fig. 3). Both these genera contain papillomaviruses that infect mammals of the order *Artiodactyla* (even-toed ungulates) including dromedaries, cervids, cows, bison, and sheep. CdPV1 and CdPV2 shared similar values of nucleotide identity of L1 with members of the genera *Deltapapillomavirus* and *Epsilonpapillomavirus*, with mean values of 64.0 and 62.2 % respectively (Table 1). In addition, pairwise comparisons of the L1 ORF of the PVs within these genera demonstrated nucleotide identities of more than 60 % (Table 2).

Both clones were deposited in the new International Reference Centre for Animal Papillomavirus at the American Museum of Natural History (New York, USA).

Discussion

Here we report for the first time the complete genomes of *Camelus dromedarius* papillomavirus types 1 and 2, tentatively classified within the genus *Deltapapillomavirus*. These are the first complete PVs characterized in species of the suborder *Tylopoda*. All other artiodactyl PVs described so far are from members of the suborder *Ruminantia*.

Both viruses were isolated from fibropapillomas, although no putative E5 proteins homologous to E5 of BPV1 were identified. The BPV1 E5 is known for its fibroblast transforming activities (Talbert-Slagle & DiMaio, 2009), and the absence of BPV1-like E5 proteins in CdPV1 and CdPV2 raises questions about alternative E5-like proteins and their possible role in fibroblast transformation. Hitherto, all PVs within the genera

Deltapillomavirus and *Epsilonpapillomavirus* have been described as harbouring E5 or at least a hydrophobic putative protein between the E2-L2 region (Bravo & Alonso, 2004; Erdelyi *et al.*, 2008; Tomita *et al.*, 2007). Since we identified short ORFs between early and late genes that might represent E5-like proteins it would be interesting to express these proteins and analyze their fibroblast transforming capability.

The primary target of the papillomavirus E7 protein is the retinoblastoma family proteins (pRb), whose inactivation is important for the differentiation-dependent productive viral lifecycle and for tumour progression (Moody & Laimins, 2010). However, none of the E7 proteins of CdPV1 or CdPV2 demonstrated the defined pRb binding domain, as was also the case with the E7 proteins of the delta- and epsilonpapillomaviruses (Erdelyi *et al.*, 2008; Narechania *et al.*, 2004; Tomita *et al.*, 2007). The lack of the pRB binding domain in E7 proteins of PVs within these genera has been suggested to be a marker for the development of fibropapillomas, which may not be a characteristic that is only attributable to E5 (Narechania *et al.*, 2004).

The L1 ORF is the most conserved region in the papillomavirus, and has been used successfully for phylogeny and taxonomy (Bernard, 2005; Bernard *et al.*, 2010). We grouped CdPV1 and 2 within the deltapapillomaviruses due to a combination of facts, namely that their average L1-nucleotide identity was highest to PVs in that genus (Table 1) and that their position in the phylogenetic tree was slightly closer to the deltapapillomavirus than to the epsilonpapapillomavirus (Fig. 3). According to the current classification system, PVs of a genus share more than 60 % nucleotide identity in the L1 ORF (Bernard *et al.*, 2010; de Villiers *et al.*, 2004). In the present study, the pairwise nucleotide comparisons of the members of the epsilon and deltapapillomaviruses showed values above the range of 60 % (Table 2), indicating that all these PVs can be lumped into one genus. Due to the relatively close distance

between these genera it has also been suggested that the delta- and epsilonpapillomaviruses should be classified within one genus if using an alternative classification system based on the E1–E2 ORFs (Bravo & Alonso, 2007) as well as a system of amino acid sequences of compiled ORFs (E6-E7-E2-E1-L2-L1) (Narechania *et al.*, 2004). Furthermore, the International Committee for the Taxonomy of Viruses (ICTV) recommends inclusion of both genotypic and phenotypic information in the definition of viral genera and species, i.e. information on the ecological niche and the relational properties of viral and host components (Bernard *et al.*, 2010; van Regenmortel, 2000). Concerning the delta- and epsilonpapillomaviruses, they manifest L1 ORF nucleotide identities higher than the percentage-border of a genus, but also exclusively infect ungulates, and are the etiological agents of fibropapillomas. Thus, the available data of nucleotide similarity and tropism of these PVs strongly suggest that it might be worth considering merging these two genera into one genus, tentatively labelled as delta-epsilonpapillomavirus, with the inclusion of the dromedary papillomaviruses.

A recent report showed a corneal papilloma from an adult camel (15 years), in which BPV1 was suggested as the possible cause of the disease (Kilic *et al.*, 2010). In that study, the presence of a virus was shown by immunohistochemistry using an antibody to BPV1. The viral antigen was observed in keratinocytes as well as in fibrocytes. Speculatively, the corneal papilloma may have contained CdPV that cross-reacted with the BPV1-antibodies, which is supported by our observation of the relatively close relationships between the CdPVs and BPV1. Furthermore, five fibropapillomas from llamas and alpacas have been reported (Schulman *et al.*, 2003). From one of these lesions an amplicon (ca. 160 bp) was sequenced that was most closely related to BPV1 (73 %). The corresponding identities to CdPV1 and CdPV2 were 76 % and 72 %,

respectively (data not shown), indicating that additional PV types within the genus Delta exist.

An interesting feature of the outbreak in our study was that only young animals (<8 months old) were affected. This is in agreement with the survey of camel papillomatosis carried out between 1992 and 1994 in eastern Sudan by Khalafalla *et al.* (1998). That work revealed a 3.3 % morbidity rate predominantly in young animals 3-14 months of age. This suggests that the immune system of adults is capable of controlling these papillomaviruses and that CdPV1 and 2 are immunogenic. Nevertheless, in the present study the cumulative incidence was about 20 % for the herd and 44 % for the young animals. The higher morbidity observed here could be due to differences in the number of young animals studied or the time of the year in which the animals were examined, since not only papillomatosis, but also camel contagious ecthyma and camel pox have been associated with the rainy season (Khalafalla, 1998). Also, the time point for examination of animals could influence the prevalence of papillomatosis since it has been reported that within three weeks virtually all young animals in a herd are infected (Munz *et al.*, 1990). Even though papillomatosis is a mild disease in dromedaries, and recurrence has not been reported, the high morbidity shown in this and previous studies highlights the importance of understanding the viral genomic sequences for surveillance of the disease, as new outbreaks with a higher severity might occur.

A limitation of the study was that only two samples were taken from the herd. This was done in order to reduce stress to the animals. As a consequence, the question remains of whether the characterized viruses were also present in lesions from other animals in the herd. It is noteworthy that two different PVs were identified from the only two lesions analyzed, and they were distant enough to be considered as two different species. It is likely that a wide diversity of PVs exists in dromedaries since several PVs have been

10

described in other animal species. For example, 10 PV types have been identified in domestic cows and about 120 HPVs are characterized (Bernard *et al.*, 2010). Moreover, the diversity seems to be even broader as about 130 putative types have also been observed on the human skin (Forslund, 2007).

In conclusion, we showed for the first time the genomic characteristics of papillomaviruses infecting a camelid. The genetic information will be useful for evolutionary studies of the family *Papillomaviridae* as well as for the development of diagnostic methods for surveillance of the disease in camels.

Materials and methods

Sample collection

All animals in the farm (n = 55) were restrained and the skin all over the body was inspected for abnormal lesions. Animals with skin lesions were identified, and the lesions were described and photographed.

Whole biopsies from affected animals were excised surgically (lidocaine 2 % was used as a local anaesthetic) and fixed in 10 % neutral buffered formalin. Samples for DNA extraction were preserved in 50 % buffered glycerol.

Histopathology

Samples for histopathology were trimmed, serially dehydrated, paraffin embedded, sectioned at 5 µm and stained with hematoxylin and eosin.

DNA extraction, cloning and sequencing

DNA was extracted from two biopsies of about 3 mm diameter, by incubation overnight at 37°C, with 400 µl `lysis buffer' (10 mM Tris-HCl, 10 mM NaCl, 10 mM EDTA, pH 7.8) with 4 % SDS and proteinase K (200 µg/ml) (Boehringer Mannheim), and then 200 µl was processed through the automated MagNA Pure LC using the Total Nucleic acid kit (Roche) and eluted in 100 µl. From a dilution 1:10, 5 µl was used for amplification of the papillomavirus genome by PCR using FA-primers targeting the L1 ORF (Forslund *et al.*, 1999). These PCR products were sequenced and used to design primers with Oligo 7 software (Molecular Biology Insights, USA) that were used to amplify the complete genomes with Expand Long Template PCR kit (Roche). The reaction was prompted in buffer 2 according to the manufacturer's instructions. The following primers were used: CdPV1 FAP F (CTACCTGCCGGATCATGTCCA), CdPV1 FAP R

(ATCAGCTCCTGCACATGCTT), CdPV2 FAP F (CAATTTAGAGTGTTCAAAGTGCAA) and CdPV2 FAP R (ATTGGGGGATACTTTTGGTATGT). Forty-four cycles of PCR were performed with the following conditions: 2 min at 94 ℃; 10 cycles of 10 sec at 94 ℃, 30 sec at 68 ℃ with a decrease of 0.5 ℃ per cycle and 6 min at 68 ℃; 10 cycles of 10 sec at 94 ℃, 30 sec at 59 ℃ and 6 min at 68 ℃; 24 cycles of 15 sec at 94 ℃, 30 sec at 57 ℃ and 6 min at 68 ℃ with an increase of 10 sec per cycle; and finally 10 min at 68 ℃. PCR products were visualized by 0.6 % agarose gel electrophoresis, and the bands of about 8 kb were excised from the gel without UV irradiation (the position of the bands was assessed with an aliquot run in a parallel lane). The product was further purified with QIAquick Gel Extraction Kit (QIAGEN) and cloned using TOPO-TA cloning kit (Invitrogen), according to the manufacturer's instructions. Transformed bacteria were identified by colony PCR with FAP59 and M13 (-29 and -43) primers. The plasmids were purified with QIAprep Spin Miniprep kit (QIAGEN) and sequenced by primer walking (ABI 3730XL, MWG, Germany).

Bioinformatic analysis

Sequence handling, including pattern searches (Smith-Waterman algorithm) to identify common papillomavirus features as well as annotation of the viral genome for GenBank were carried out using the UGENE software v1.8.0 (http://genome.unipro.ru). Putative proteins from ORFs were generated by UGENE and they were then searched for similarities with other proteins using BLASTp (http://blast.ncbi.nlm.nih.gov). Proteins were analyzed for unique domains with ScanProsite (http://www.expasy.ch/prosite) and SMART (Letunic *et al.*, 2009), including searches in the Pfam database (http://pfam.sanger.ac.se/search) (Finn *et al.*, 2010). Hydrophobicity profiles were

calculated with the online tool ProtScale (Gasteiger *et al.*, 2005) (http://expasy.org/cgi-bin/protscale.pl), using the Kyte & Doolittle scale (Kyte & Doolittle, 1982) and a window size of 13 aa. Transmembrane segments were predicted with the TMHMM algorithm on the HUSAR server (http://genome.dkfz-heidelberg.de). The grand average of hydropathy (GRAVY), defined by the sum of hydropathy values of all amino acids divided by the protein length, was calculated with the GRAVY Calculator (http://www.gravy-calculator.de). Pairwise identity values from nucleotides and proteins were calculated using the MatGAT program (Campanella *et al.*, 2003).

Phylogenetic inference

L1 nucleotide sequences (including stop codons) of CdPV1, CdPV2 and of 71 PV types, each representing the prototype for each species according to ICTV (Bernard *et al.*, 2010; Fauquet *et al.*, 2005), and recently reported genomes were used for phylogenetic analysis (see supplementary table 4 for a list of accession numbers).

A multiple alignment was made using Muscle version 3.8 (Edgar, 2004) plug-in from the UGENE package. With the software jModelTest (Guindon & Gascuel, 2003; Posada, 2008), we examined 88 models of nucleotide substitution for our data. Setting the Akaike Information Criterion, jModelTest selected the General Time Reversible as the best-fitting model, with a rate variation showing a gamma distribution and a proportion of invariant sites (GTR+Γ+I). We made a Bayesian phylogenetic inference with MrBayes 3.2 (Ronquist & Huelsenbeck, 2003) using the GTR+Γ+I model (settings: nst = 6 nucmodel = 4by4 rates = invgamma). Six independent analyses, each with four Markov chains (one cold), were run for 10 000 000 generations, sampling every 100 generations. Prior to termination of the run, we analyzed the posterior sample of trees using the diagnostic software AWTY (Nylander *et al.*, 2008). Parameters were

also checked for acceptable mixing and convergence to the stationary distribution with the program Tracer 1.5 (Rambaut & Drummond, 2009).

Trees generated before the stabilization of the likelihood scores were discarded (25 % burn in). We used all six runs to generate a consensus tree with clade-credibility values (posterior probabilities) and branch lengths and used the program FigTree 1.3.1 (Rambaut, 2009) to make a graphical representation.

References

Alberti, A., Pirino, S., Pintore, F., Addis, M. F., Chessa, B., Cacciotto, C., Cubeddu, T., Anfossi, A., Benenati, G. & other authors (2010). Ovis aries Papillomavirus 3: A prototype of a novel genus in the family Papillomaviridae associated with ovine squamous cell carcinoma. *Virology* **407**, 352-359.

Androphy, E. J., Lowy, D. R. & Schiller, J. T. (1987). Bovine papillomavirus E2 trans-activating gene product binds to specific sites in papillomavirus DNA. *Nature* **325**, 70-73.

Barbosa, M. S., Lowy, D. R. & Schiller, J. T. (1989). Papillomavirus polypeptides E6 and E7 are zinc-binding proteins. *J Virol* **63**, 1404-1407.

Bernard, H. U. (2005). The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *J Clin Virol* **32 Suppl 1**, S1-6.

Bernard, H. U., Burk, R. D., Chen, Z., van Doorslaer, K., Hausen, H. & de Villiers, E. M. (2010). Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* **401**, 70-79.

Bravo, I. G. & Alonso, A. (2004). Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *J Virol* **78**, 13613-13626.

Bravo, I. G. & Alonso, A. (2007). Phylogeny and evolution of papillomaviruses based on the E1 and E2 proteins. *Virus Genes* **34**, 249-262.

Campanella, J. J., Bitincka, L. & Smalley, J. (2003). MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* **4**, 29.

Chen, G. & Stenlund, A. (2001). The E1 initiator recognizes multiple overlapping sites in the papillomavirus origin of DNA replication. *J Virol* **75**, 292-302.

Chow, L. T., Broker, T. R. & Steinberg, B. M. (2010). The natural history of human papillomavirus infections of the mucosal epithelia. *APMIS* **118**, 422-449.

Dahiya, A., Gavin, M. R., Luo, R. X. & Dean, D. C. (2000). Role of the LXCXE binding site in Rb function. *Mol Cell Biol* **20**, 6799-6805.

de Villiers, E. M., Fauquet, C., Broker, T. R., Bernard, H. U. & zur Hausen, H. (2004). Classification of papillomaviruses. *Virology* **324**, 17-27.

Doorbar, J. & Myers, G. (1996). The E4 Protein. In *Human Papillomaviruses 1996 Compendium*, vol. III, pp. 58-80.

Doorbar, J., Campbell, D., Grand, R. J. & Gallimore, P. H. (1986). Identification of the human papilloma virus-1a E4 gene products. *EMBO J* **5**, 355-362.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

Erdelyi, K., Balint, A., Dencso, L., Dan, A. & Ursu, K. (2008). Characterisation of the first complete genome sequence of the roe deer (Capreolus capreolus) papillomavirus. *Virus Res* **135**, 307-311.

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. & Ball, L. A. (2005). *Eigth Report of the International Committee on Taxonomy of Viruses*. London: Elsevier Academic Press.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G. & other authors (2010). The Pfam protein families database. *Nucleic Acids Res* **38**, D211-222.

Forslund, O. (2007). Genetic diversity of cutaneous human papillomaviruses. *J Gen Virol* **88**, 2662-2669.

Forslund, O., Antonsson, A., Nordin, P., Stenquist, B. & Hansson, B. G. (1999). A broad range of human papillomavirus types detected with a general PCR method suitable for analysis of cutaneous tumours and normal skin. *J Gen Virol* **80**, 2437-2443.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, pp. 571-607. Edited by J. M. Walker: Humana Press.

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704.

Hegde, R. S. (2002). The papillomavirus E2 proteins: structure, function, and biology. *Annu Rev Biophys Biomol Struct* **31**, 343-360.

Hegde, R. S., Grossman, S. R., Laimins, L. A. & Sigler, P. B. (1992). Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature* **359**, 505-512.

Iyer, L. M., Leipe, D. D., Koonin, E. V. & Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *J Struct Biol* **146**, 11-31.

Khalafalla, A. I. (1998). Ecthyma and Camel Papillomatosis in the Sudan. In *Proceedings of the Third Annual Meeting for Animal Production Under Arid Conditions*, pp. 151-131.

Khalafalla, A. I., Abbas, Z. & Mohamed, M. E. H. (1998). Camel papillomatosis in the Sudan. *Journal of camel Practice and Research* **5**, 157-159.

Kilic, N., Toplu, N., Aydogan, A., Yaygingul, R. & Ozsoy, S. Y. (2010). Corneal papilloma associated with papillomavirus in a one-humped camel (Camelus dromedarius). *Vet Ophthalmol* **13 Suppl**, 100-102.

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132.

Letunic, I., Doerks, T. & Bork, P. (2009). SMART 6: recent updates and new developments. *Nucleic Acids Res* **37**, D229-232.

Li, R., Knight, J., Bream, G., Stenlund, A. & Botchan, M. (1989). Specific recognition nucleotides and their DNA context determine the affinity of E2 protein for 17 binding sites in the BPV-1 genome. *Genes Dev* **3**, 510-526.

Liu, X., Schuck, S. & Stenlund, A. (2010). Structure-based mutational analysis of the bovine papillomavirus E1 helicase domain identifies residues involved in the nonspecific DNA binding activity required for double trimer formation. *J Virol* **84**, 4264-4276.

Moody, C. A. & Laimins, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* **10**, 550-560.

Munz, E., Moallin, A. S., Mahnel, H. & Reimann, M. (1990). Camel papillomatosis in Somalia. *Zentralbl Veterinarmed B* **37**, 191-196.

Narechania, A., Terai, M., Chen, Z., DeSalle, R. & Burk, R. D. (2004). Lack of the canonical pRB-binding domain in the E7 ORF of artiodactyl papillomaviruses is associated with the development of fibropapillomas. *J Gen Virol* **85**, 1243-1250.

Nylander, J. A., Wilgenbusch, J. C., Warren, D. L. & Swofford, D. L. (2008). AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**, 581-583.

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**, 1253-1256.

Rambaut, A. (2009). FigTree v1.3. Available from: http://beast.bio.ed.ac.uk/figtree.

Rambaut, A. & Drummond, A. J. (2009). Tracer v1.5. Available from: http://beast.bio.ed.ac.uk/Tracer.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Schulman, F. Y., Krafft, A. E., Janczewski, T., Reupert, R., Jackson, K. & Garner, M. M. (2003). Camelid mucoutaneous fibropapillomas: clinicopathologic findings and association with papillomavirus. *Vet Pathol* **40**, 103-107.

Spalholz, B. A., McBride, A. A., Sarafi, T. & Quintero, J. (1993). Binding of bovine papillomavirus E1 to the origin is not sufficient for DNA replication. *Virology* **193**, 201-212.

Talbert-Slagle, K. & DiMaio, D. (2009). The bovine papillomavirus E5 protein and the PDGF β receptor: It takes two to tango. *Virology* **384**, 345-351.

Tomita, Y., Literak, I., Ogawa, T., Jin, Z. & Shirasawa, H. (2007). Complete genomes and phylogenetic positions of bovine papillomavirus type 8 and a variant type from a European bison. *Virus Genes* **35**, 243-249.

Ustav, E., Ustav, M., Szymanski, P. & Stenlund, A. (1993). The bovine papillomavirus origin of replication requires a binding site for the E2 transcriptional activator. *Proc Natl Acad Sci U S A* **90**, 898-902.

Ustav, M. & Stenlund, A. (1991). Transient replication of BPV-1 requires two viral polypeptides encoded by the E1 and E2 open reading frames. *EMBO J* **10**, 449-457.

van Regenmortel, M. H. V. (2000). In *Seventh Report of the International Committee on Taxonomy of Viruses*, pp. 3-16. Edited by M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle & R. B. Wikner. New York: Academic Press.

Yang, L. & Botchan, M. (1990). Replication of bovine papillomavirus type 1 DNA initiates within an E2-responsive enhancer element. *J Virol* **64**, 5903-5911.

Yang, L., Li, R., Mohr, I. J., Clark, R. & Botchan, M. R. (1991). Activation of BPV-1 replication in vitro by the transcription factor E2. *Nature* **353**, 628-632.

Yang, Y. C., Okayama, H. & Howley, P. M. (1985). Bovine papillomavirus contains multiple transforming genes. *Proc Natl Acad Sci U S A* **82**, 1030-1034.

Zhou, J., Doorbar, J., Sun, X. Y., Crawford, L. V., McLean, C. S. & Frazer, I. H. (1991). Identification of the nuclear localization signal of human papillomavirus type 16 L1 protein. *Virology* **185**, 625-632.

Figure 1
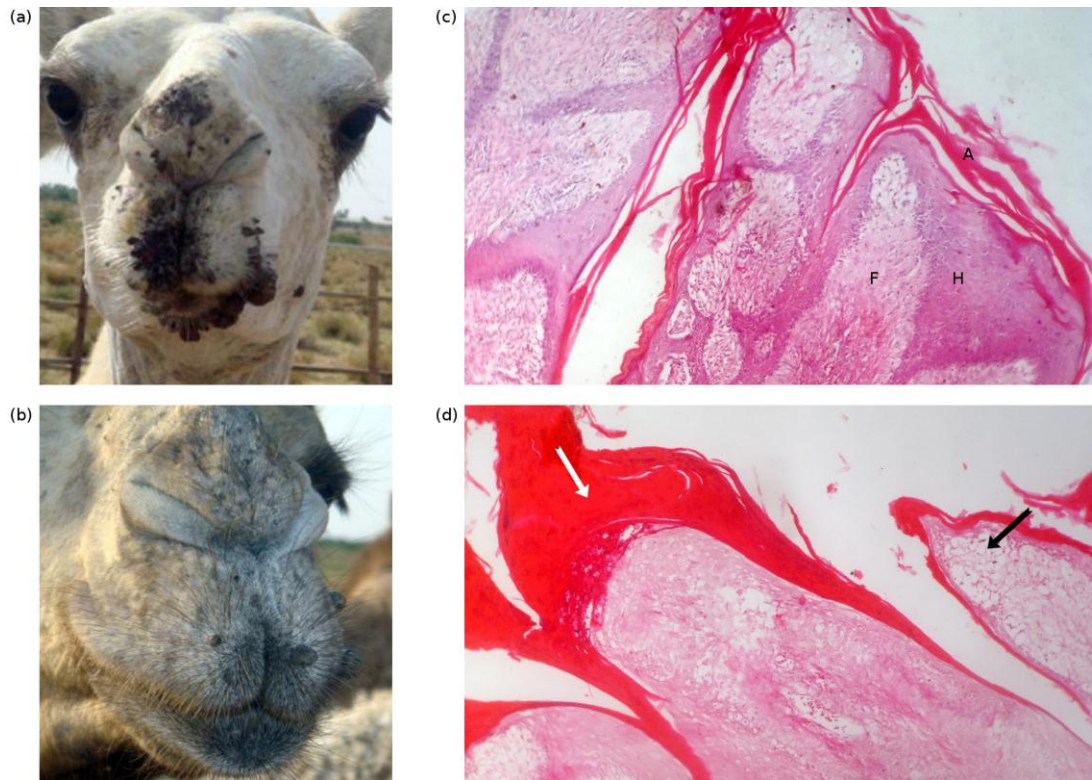
Papillomatosis in dromedary calves. (**a**) Young dromedary showing cauliflower–like papillomas in the muzzle. (**b**) Calf showing multiple round nodules in the upper lip. (**c**) Histopathological section (×100) showing hyperkeratosis, A; acanthosis, H and dermal fibrosis, F. (**d**) Histopathological section (×100) showing multiple papillary proliferations. Note hyperkeratosis (white arrow) and hydropic degeneration in keratinocytes (black arrow). Panels **a** and **c** refer to the sample from which CdPV1 was isolated, and panels **b** and **d** to the one for CdPV2.

Figure 2
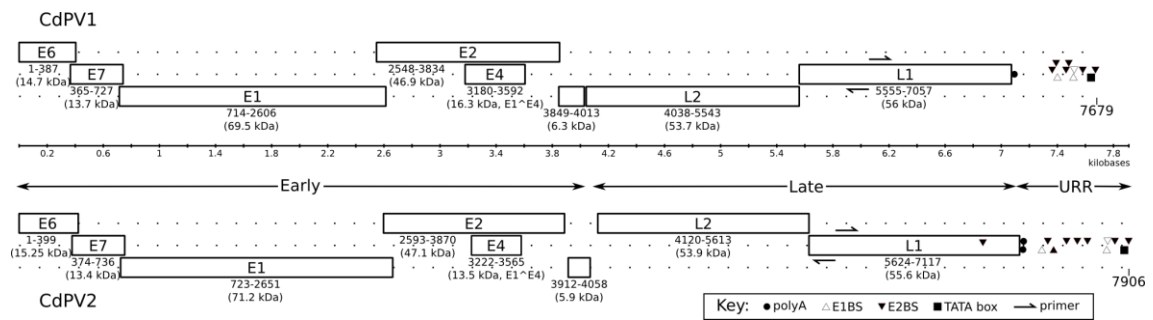
Genomic organization of CdPV1 and CdPV2. The three main regions of the genome are depicted below the ruler. Numbers below each ORF (white boxes) indicate nucleotide position from start to stop codon and the molecular mass for the corresponding putative protein. The E4 ORF does not contain a start codon (ATG), but an ATG is probably added by splicing of E4 mRNA to E1-mRNA generating an E4 protein which starts with five amino acids from E1 (E1^E4). Putative binding sites for viral proteins and cellular factors are shown as follows: E2 binding site, E2BS (▼); E1 binding site, E1BS (Δ); TATA-box (■); Polyadenylation signal (●). The position of primers used for amplification of the genome and further cloning are indicated with half arrows.

Figure 3
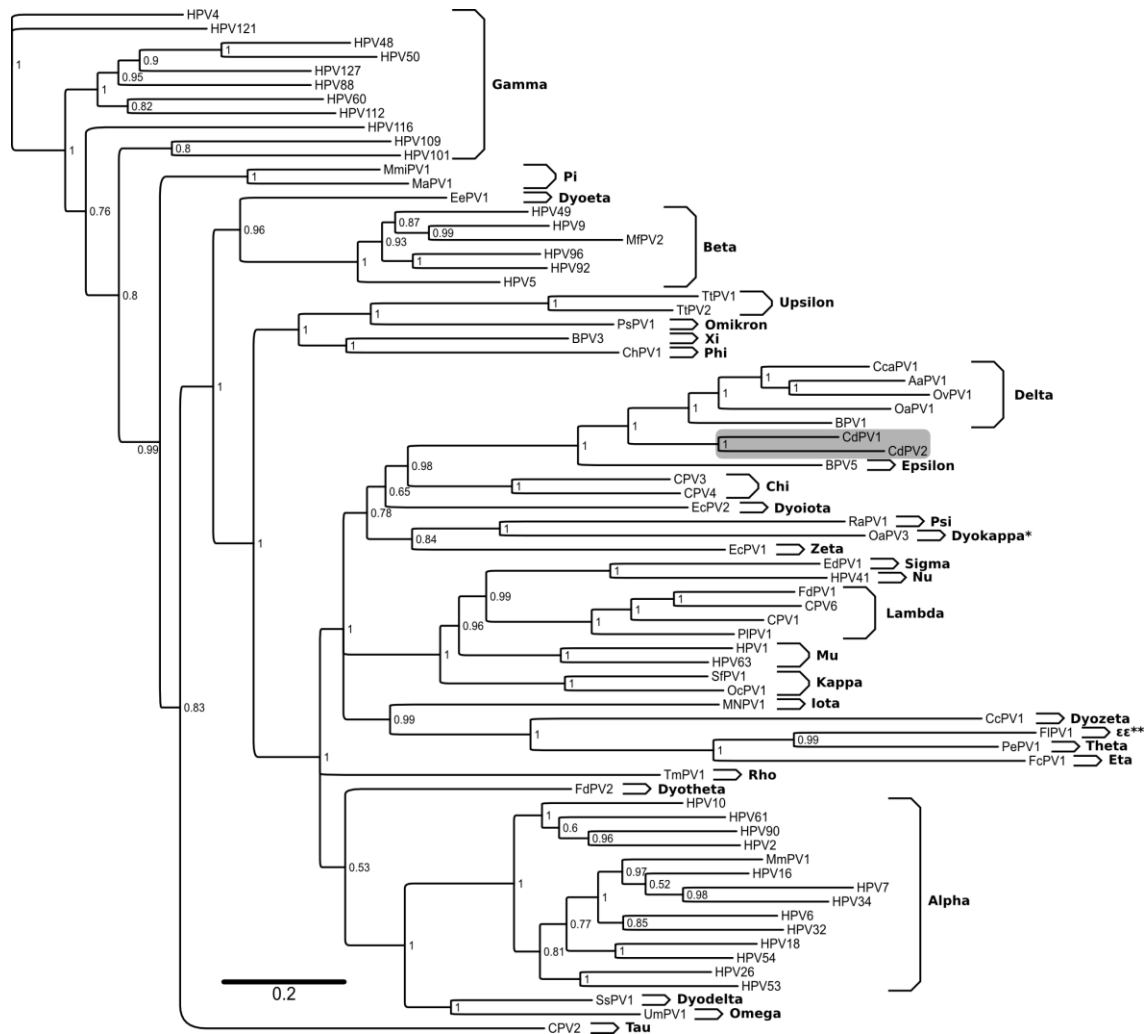
Bayesian phylogenetic tree inferred from the L1 nucleotide sequence of CdPV1 and CdPV2, and of 71 representative papillomaviruses. Acronyms are according to Bernard *et al.* (2010) and genera (Fauquet *et al.*, 2005) indicated in bold. CdPV1 and CdPV2 are located in a separate branch between delta- and epsilonpapillomaviruses. Numbers on nodes are the supporting posterior probability of the split. Branch lengths are displayed to scale according to the bar (0.2 substitutions per site). * Genus suggested by the authors of the paper describing the virus and not yet officially recognized by the ICTV. ** *Dyoepsilon*. CcaPV, *Capreolus capreolus* papillomavirus; CcPV, *Caretta caretta* papillomavirus; ChPV, *Capra hircus* papillomavirus; CPV, *Canis familiaris* papillomavirus; EcPV, *Equus caballus* papillomavirus; EdPV, *Erethizon dorsatum*

papillomavirus; EePV, *Erinaceus europaeus* papillomavirus; FcPV, *Fringilla coelebs* papillomavirus; FdPV, *Felis domesticus* papillomavirus; FlPV, *Francolinus leucoscepus* papillomavirus; HPV, human papillomavirus; MaPV, *Mesocricetus auratus* papillomavirus; MfPV, *Macaca fascicularis* papillomavirus; MmiPV, *Micromys minutus* papillomavirus 1; MmPV, *Macaca mulata* papillomavirus; MNPV, *Mastomys natalensis* papillomavirus; OcPV, *Oryctolagus cuniculus* papillomavirus; PePV, *Psittacus erithacus* papillomavirus; PlPV, *Procyon lotor* papillomavirus; PsPV, *Phocoena spinipinnis* papillomavirus; RaPV, *Rousettus aegyptiacus* papillomavirus; SfPV, *Sylvilagus floridanus* papillomavirus; TmPV, *Trichechus manatus* latirostris papillomavirus; TtPV, *Tursiops truncatus* papillomavirus; UmPV, *Ursus maritimus* papillomavirus.

Table 1

CdPV1 and CdPV2 nucleotide and amino acid identities with members of the *Delta-* and *Epsilonpapillomavirus* genera.

Virus*	Genus	E6	E7	E1	E2	E4	L2	L1	Accession No.
CdPV2	δ	52.6 (39.8)	56.9 (47.2)	66.3 (62.4)	62.1 (51.6)	52.2 (45.3)	58.7 (50.5)	69.2 (72.8)	HQ912791
BPV1	δ	53.9 (39.1)	49.0 (30.8)	60.0 (52.0)	44.6 (32.3)	49.9 (25.2)	52.7 (38.4)	66.3 (65.0)	NC_001522
BPV2	δ	51.8 (38.4)	48.9 (31.6)	59.3 (51.8)	54.4 (40.6)	46.8 (32.4)	51.3 (37.9)	65.5 (64.6)	M20219
CcaPV1	δ	52.8 (42.0)	47.8 (33.3)	59.8 (52.3)	55.4 (43.3)	43.9 (29.0)	53.9 (38.8)	65.3 (65.6)	EF680235
RtPV1	δ	51.4 (38.6)	49.2 (29.8)	60.0 (50.7)	53.2 (43.8)	48.5 (27.3)	53.1 (39.4)	64.1 (64.9)	AF443292
AaPV1	δ	51.4 (38.1)	51.1 (34.4)	59.8 (52.4)	54.2 (39.1)	43.8 (29.4)	51.9 (38.6)	64.5 (64.7)	NC_001524
OvPV1	δ	53.4 (41.3)	49.6 (29.4)	60.5 (50.8)	53.7 (41.2)	47.3 (27.1)	53.2 (37.3)	62.7 (59.3)	NC_001523
OaPV1	δ	50.3 (40.8)	47.7 (31.5)	59.0 (52.0)	52.8 (42.5)	48.4 (31.2)	48.6 (31.3)	63.2 (62.9)	NC_001789
OaPV2	δ	49.7 (40.0)	47.2 (31.7)	59.2 (51.4)	54.7 (45.4)	48.1 (33.3)	52.2 (36.1)	65.2 (64.9)	U83595
BPV5	ε	52.4 (38.7)	48.5 (29.9)	57.3 (48.6)	52.6 (37.3)	48.3 (24.3)	54.1 (38.8)	62.2 (59.6)	NC_004195
BPV8	ε	50.7 (38.5)	49.1 (30.4)	58.2 (48.1)	53.4 (40.6)	48.6 (24.3)	53.5 (40.7)	62.1 (60.6)	NC_009752
CdPV1	δ	52.6 (39.8)	56.9 (47.2)	66.3 (62.4)	62.1 (51.6)	52.2 (45.3)	58.7 (50.5)	69.2 (72.8)	HQ912790
BPV1	δ	55.7 (44.9)	51.0 (31.9)	59.6 (48.9)	46.9 (31.7)	51.2 (27.2)	50.3 (38.3)	64.6 (64.1)	NC_001522
BPV2	δ	55.8 (44.2)	47.5 (29.0)	59.1 (48.1)	58.2 (42.3)	54.9 (32.2)	52.1 (37.4)	64.6 (64.2)	M20219
CcaPV1	δ	49.4 (32.4)	50.7 (30.7)	59.9 (51.7)	55.0 (41.8)	51.4 (33.3)	53.9 (39.1)	64.0 (64.0)	EF680235
RtPV1	δ	51.4 (36.4)	44.7 (31.1)	57.9 (47.9)	55.3 (40.8)	49.6 (27.1)	53.3 (41.4)	63.3 (62.5)	AF443292
AaPV1	δ	49.8 (35.7)	48.9 (31.4)	59.6 (50.3)	54.7 (38.4)	53.2 (29.3)	53.0 (41.1)	63.0 (62.1)	NC_001524
OvPV1	δ	48.8 (35.5)	50.4 (30.5)	58.2 (48.8)	54.2 (39.2)	48.7 (29.9)	51.6 (37.0)	61.4 (58.8)	NC_001523
OaPV1	δ	51.0 (41.9)	46.7 (29.0)	59.4 (52.8)	52.5 (41.2)	54.4 (34.2)	49.2 (34.7)	62.6 (62.3)	NC_001789
OaPV2	δ	50.6 (39.7)	48.6 (28.2)	59.9 (52.5)	55.1 (42.6)	55.1 (34.7)	51.7 (41.1)	62.8 (63.7)	U83595
BPV5	ε	56.2 (39.4)	49.0 (29.3)	58.4 (47.3)	53.9 (39.3)	48.3 (22.6)	53.4 (40.1)	62.5 (59.8)	NC_004195
BPV8	ε	55.6 (40.1)	50.0 (27.8)	58.0 (49.7)	54.2 (38.9)	49.1 (24.0)	54.2 (40.1)	62.0 (60.0)	NC_009752

*CcaPV1, *Capreolus capreolus* papillomavirus 1; RtPV1, *Rangifer tarandus* papillomavirus 1 [Reindeer papillomavirus (RPV)]; AaPV1,

Alces alces papillomavirus 1 [European elk papillomavirus (EEPV)]; OvPV1, *Odocoileus virginianus* papillomavirus 1 [deer

papillomavirus (DPV)]; OaPV, *Ovis aries* papillomavirus.

Table 2

Pairwise nucleotide identities of L1 ORF between all members of the *Delta-* and *Epsilonpapillomavirus* genera.

Genus	Host	Acronym	CdPV2	OvPV1	BPV1	BPV2	BPV5	BPV8	AaPV1	OaPV2	OaPV1	RtPV1	CcaPV1
δ	Camelus dromedarius	CdPV1	69.2	62.7	66.3	65.5	62.2	62.1	64.5	65.2	63.2	64.0	65.3
δ	Camelus dromedarius	CdPV2		61.6	64.6	64.6	62.5	62.0	62.9	62.8	62.8	63.3	64.0
δ	Odocoileus virginianus	OvPV1 (DPV)			65.8	65.1	60.5	60.0	70.5	65.3	64.9	71.7	70.9
δ	Bos taurus	BPV1				84.4	62.4	62.8	68.3	68.7	67.9	68.6	67.7
δ	Bos taurus	BPV2					63.3	62.2	67.5	68.4	67.3	68.1	68.2
ε	Bos taurus	BPV5						74.5	62.2	62.8	62.0	63.2	61.7
ε	Bos taurus	BPV8							61.7	62.8	61.0	60.9	62.0
δ	Alces alces	AaPV1 (EEPV)								68.7	67.8	76.5	71.5
δ	Ovis aries	OaPV2									79.5	69.4	69.0
δ	Ovis aries	OaPV1										68.7	68.0
δ	Rangifer tarandus	RtPV1											70.7