JVI Journals.ASM.org

The Genome of a Tortoise Herpesvirus (Testudinid Herpesvirus 3) Has a Novel Structure and Contains a Large Region That Is Not Required for Replication *In Vitro* or Virulence *In Vivo*

Frédéric Gandar,[a,b] Gavin S. Wilkie,[c] Derek Gatherer,[d] Karen Kerr,[c] Didier Marlier,[b] Marianne Diez,[e] Rachel E. Marschang,[f] Jan Mast,[g] Benjamin G. Dewals,[a] Andrew J. Davison,[c] Alain F. C. Vanderplasschen[a]

Immunology-Vaccinology, Department of Infectious and Parasitic Diseases, Fundamental and Applied Research for Animals & Health (FARAH), Faculty of Veterinary Medicine, University of Liège, Liège, Belgium[a]; Clinic for Birds, Rabbits and Rodents, Department of Clinical Sciences, FARAH, Faculty of Veterinary Medicine, University of Liège, Liège, Belgium[b]; MRC—University of Glasgow Centre for Virus Research, Glasgow, United Kingdom[c]; Division of Biomedical and Life Sciences, Faculty of Health and Medicine, Lancaster University, Lancaster, United Kingdom[d]; Nutrition of Companion Animals, Department of Animal Production, FARAH, Faculty of Veterinary Medicine, University of Liège, Liège, Belgium[e]; Laboratory for Clinical Diagnostics, Laboklin GmbH & Co. KG, Bad Kissingen, Germany[f]; Biocontrol Department, Research Unit Electron Microscopy, Veterinary and Agrochemical Research Centre, VAR-CODA-CERVA, Ukkel, Belgium[g]

ABSTRACT

Testudinid herpesvirus 3 (TeHV-3) is the causative agent of a lethal disease affecting several tortoise species. The threat that this virus poses to endangered animals is focusing efforts on characterizing its properties, in order to enable the development of prophylactic methods. We have sequenced the genomes of the two most studied TeHV-3 strains (1976 and 4295). TeHV-3 strain 1976 has a novel genome structure and is most closely related to a turtle herpesvirus, thus supporting its classification into genus *Scutavirus*, subfamily *Alphaherpesvirinae*, family *Herpesviridae*. The sequence of strain 1976 also revealed viral counterparts of cellular interleukin-10 and semaphorin, which have not been described previously in members of subfamily *Alphaherpesvirinae*. TeHV-3 strain 4295 is a mixture of three forms (m1, m2, and M), in which, in comparison to strain 1976, the genomes exhibit large, partially overlapping deletions of 12.5 to 22.4 kb. Viral subclones representing these forms were isolated by limiting dilution assays, and each replicated in cell culture comparably to strain 1976. With the goal of testing the potential of the three forms as attenuated vaccine candidates, strain 4295 was inoculated intranasally into Hermann's tortoises (*Testudo hermanni*). All inoculated subjects died, and PCR analyses demonstrated the ability of the m2 and M forms to spread and invade the brain. In contrast, the m1 form was detected in none of the organs tested, suggesting its potential as the basis of an attenuated vaccine candidate. Our findings represent a major step toward characterizing TeHV-3 and developing prophylactic methods against it.

IMPORTANCE

Testudinid herpesvirus 3 (TeHV-3) causes a lethal disease in tortoises, several species of which are endangered. We have characterized the viral genome and used this information to take steps toward developing an attenuated vaccine. We have sequenced the genomes of two strains (1976 and 4295), compared their growth *in vitro*, and investigated the pathogenesis of strain 4295, which consists of three deletion mutants. The major findings are that (i) TeHV-3 has a novel genome structure, (ii) its closest relative is a turtle herpesvirus, (iii) it contains interleukin-10 and semaphorin genes (the first time these have been reported in an alphaherpesvirus), (iv) a sizeable region of the genome is not required for viral replication *in vitro* or virulence *in vivo*, and (v) one of the components of strain 4295, which has a deletion of 22.4 kb, exhibits properties indicating that it may serve as the starting point for an attenuated vaccine.

The order *Herpesvirales* contains a large number of enveloped, double-stranded DNA viruses that share structural, genetic, and biological properties and is divided into three families infecting a wide range of hosts (1). One of these families, the *Herpesviridae*, contains viruses infecting mammals, birds, or reptiles and is subdivided into three subfamilies, the *Alpha-*, *Beta-*, and *Gammaherpesvirinae* (2). Members of these subfamilies are referred to colloquially as alpha-, beta-, and gammaherpesviruses, respectively.

All herpesviruses of reptiles identified to date group among the alphaherpesviruses, in lineages distinct from herpesviruses of mammals or birds (3). Many of the hosts of these viruses belong to the order Testudines (also called Chelonii) and include pond turtles, marine turtles, and terrestrial tortoises. Among members of the order Testudines, herpesvirus infections have been described chiefly in the last two groups. Among marine turtles, the genome of chelonid herpesvirus 5 (ChHV-5), which is thought to be the causative agent of fibropapillomatosis, has been cloned as a bacterial artificial chromosome from infected tissue and sequenced. This virus has been classified into the genus *Scutavirus* (4). The

Received 16 July 2015 **Accepted** 27 August 2015

Accepted manuscript posted online 2 September 2015

Citation Gandar F, Wilkie GS, Gatherer D, Kerr K, Marlier D, Diez M, Marschang RE, Mast J, Dewals BG, Davison AJ, Vanderplasschen AFC. 2015. The genome of a tortoise herpesvirus (testudinid herpesvirus 3) has a novel structure and contains a large region that is not required for replication *in vitro* or virulence *in vivo*. J Virol 89:11438–11456. doi:10.1128/JVI.01794-15.

Editor: R. M. Longnecker

Address correspondence to Andrew J. Davison, andrew.davison@glasgow.ac.uk, or Alain F. C. Vanderplasschen, a.vdplasschen@ulg.ac.be.

A.J.D. and A.F.C.V. contributed equally to this article.

phylogenetic relationship between herpesviruses infecting marine turtles and those infecting tortoises is unclear (3).

Tortoises exist as at least 40 species belonging to the family Testudinidae. Herpesviruses have been isolated from healthy or sick individuals belonging to several of these species. Based on partial sequencing of the viral DNA polymerase gene, four genotypes have been identified, leading to the nomenclature testudinid herpesvirus 1 to 4 (TeHV-1 to TeHV-4) (3). Among these genotypes, TeHV-3 appears to be the most pathogenic and has been shown to affect several tortoise species, with those from the genus *Testudo* (e.g., *Testudo hermanni*) being the most sensitive (5, 6). Young tortoises are more susceptible to TeHV-3 disease than adults and can suffer from mortality rates of up to 100%. These pathological and epidemiological features and the fact that many of the susceptible host species are endangered contribute to ecological concerns over this virus. Clinical signs depend on several factors, including host species, age, the season at which infection occurs, and the viral strain involved (7–12). The main clinical signs are nasal discharge, rhinitis, conjunctivitis associated with blepharospasm, and diphtheritic plaques in the oral cavity and esophagus (13). Weight loss, cachexia, central nervous symptoms (such as circling and head tilt), and death are observed in advanced stages of the disease. The virus has been isolated from several tissues, suggesting a broad tropism (8, 11, 14, 15). With the goal of controlling the threat that TeHV-3 poses to tortoise populations, various inactivated vaccine candidates have been tested, but none has proved efficacious (16, 17). Obvious alternatives to inactivated vaccines are attenuated vaccines and subunit vaccines targeting key viral proteins. The knowledge required for the development of such vaccines would include the genome sequence.

Here, we sequenced the genomes of representative TeHV-3 strains 1976 (18) and 4295 (9). Strain 1976 has a genome structure not reported previously among herpesviruses and is most closely related phylogenetically to turtle herpesviruses. This strain also contains genes with cellular homologs that have not been described previously in alphaherpesviruses. Strain 4295 consists of a mixture of three forms, of which the genomes exhibit large, partially overlapping deletions in comparison with strain 1976. The effects of these deletions on viral growth *in vitro* and virulence *in vivo* were investigated.

MATERIALS AND METHODS

Cells and viruses. Terrapene heart cells (TH-1, subline B1) (19) were cultured in Dulbecco's modified Eagle's medium (DMEM; Sigma-Aldrich) containing 4.5 g/liter glucose, 5% fetal calf serum, and 1% nonessential amino acids (Invitrogen). Cells were cultured at 25°C in a humid atmosphere in the presence of 5% CO_2. Two previously described TeHV-3 strains were used. Strain 1976 (passage 6) originated from the intestine of a Horsfield's tortoise (*Testudo horsfieldii*) that died from TeHV-3 infection (18). Strain 4295 (passage 14) was isolated from a pharyngeal swab performed on a clinically healthy Hermann's tortoise (*T. hermanni*) during an outbreak of herpesviral disease (9). The absence of extraneous contaminating agents in the TH-1 cells and TeHV-3 strains was confirmed by electron microscopic examination of mock-infected and infected TH-1 cells (Fig. 1). Clones of strain 4295 were produced by limiting dilution. Tenfold serial dilutions of infected culture supernatant were inoculated onto TH-1 cells grown in 96-well plates. Clones were amplified from dilutions for which less than 10% of the wells showed signs of infection. The purity and genotype of the clones were determined by PCR (see below; see also Fig. 7A and B).

Virion DNA production. Confluent TH-1 cells were infected with TeHV-3 at a multiplicity of infection (MOI) of 0.2 PFU/cell. To reduce contamination by cellular DNA, cell supernatant was harvested at the early stages of viral release, when approximately 10% of cells were exhibiting cytopathic effect. Virions were semipurified as described previously (20). Briefly, after removal of the cell debris by centrifugation (1,000 × *g*, 10 min, 4°C), viral particles were pelleted by ultracentrifugation through a 30% sucrose cushion (100,000 × *g*, 2 h, 4°C). DNA was purified from virions as described previously (21).

DNA sequencing. Virion DNA (1 μg) was sheared by sonication to an average size of 470 bp, and sequencing libraries were prepared by using a KAPA library preparation kit (KAPA Biosystems). The fragments were A-tailed and ligated to the NEBnext Illumina adaptor (New England BioLabs), and NEBnext indexing primers were added by carrying out four cycles of PCR in an ABI 7500 real-time cycler, using a KAPA HiFi Real-time library amplification kit. The libraries were sequenced by using a MiSeq (Illumina) operating v2 chemistry, generating data sets of 250-nucleotide (nt) paired-end reads. The reads were filtered for quality, adapter sequences were removed by using Trim Galore v. 0.2.2 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore), and the reads were assembled by using Velvet v. 1.2.07 (22) and AbySS v. 1.3.5 (23). Larger contigs were generated from the assemblies by using Phrap v. 1.080812 (24, 25) and then IMAGE v. 2.31 (26). The data sets were assembled against the contigs by using BWA v. 0.6.2-r126 (27), and the alignments were inspected by using Tablet 1.13.12.17 (28). Problematic regions, including those representing deletions or duplications, and those containing relatively short reiterated sequences (including mononucleotide tracts) were resolved by using custom Perl scripts to count and extract individual reads from the data sets for further analysis (29). The sizes of the more substantial reiterated sequences were estimated by PCR (Table 1), and in most instances, the products were sequenced by using Sanger technology.

Potential genome termini were identified in the BWA alignments from sets of reads sharing a common end. They were confirmed for strains 1976 and 4295 (prior to subcloning) by using a published methodology (30) that involves ligating partially double-stranded DNA adapters to blunt-ended or untreated viral DNA, followed by PCR using a combination of a virus-specific primer and an adapter-specific primer (the former set of primers are listed in Table 1). The inserts from 12 plasmids generated from each purified PCR product were sequenced by using Sanger technology, and the genome termini were defined as being located at the positions represented by the majority of clones. Analysis of both blunt-ended and untreated viral DNA allowed unpaired nucleotides at the 3′-ends of the genome to be identified. The genome termini of the three subclones (m1, m2, and M) of strain 4295 were not determined in this way but rather inferred from the data from strain 4295. The final genome sequences of strain 1976 and the three subclones (m1, m2, and M) of strain 4295 were constructed on the basis of the locations of the termini, and the integrity of each was verified by aligning it against the relevant data set using BWA, visualizing the alignment by using Tablet. The sequences were deposited in NCBI GenBank (see below).

Southern blot analysis. Southern blot analysis of virion DNA, digested with EcoRI or KpnI restriction endonucleases, was performed as described previously (31, 32). Probes were produced by PCR amplification of strain 1976 DNA using specific primers (Table 1).

Viral growth curves. Triplicate cultures of TH-1 cells in 24-well plates were inoculated with TeHV-3 at an MOI of 0.2 PFU/cell. After an incubation period of 4 h, the cells were washed with phosphate-buffered saline (PBS) and overlaid with DMEM containing 4.5 g/liter glucose, 5% fetal calf serum (FCS), and 1% nonessential amino acids (Invitrogen). The supernatants of infected cultures were harvested at successive intervals after infection and stored at −80°C. The amount of infectious virus was determined by plaque assay on TH-1 cells as described previously (32, 33). The data, expressed as mean titers and standard deviations (SD) of triplicate assays, were analyzed for significant differences ($P < 0.05$) using one-way analysis of variance (ANOVA).

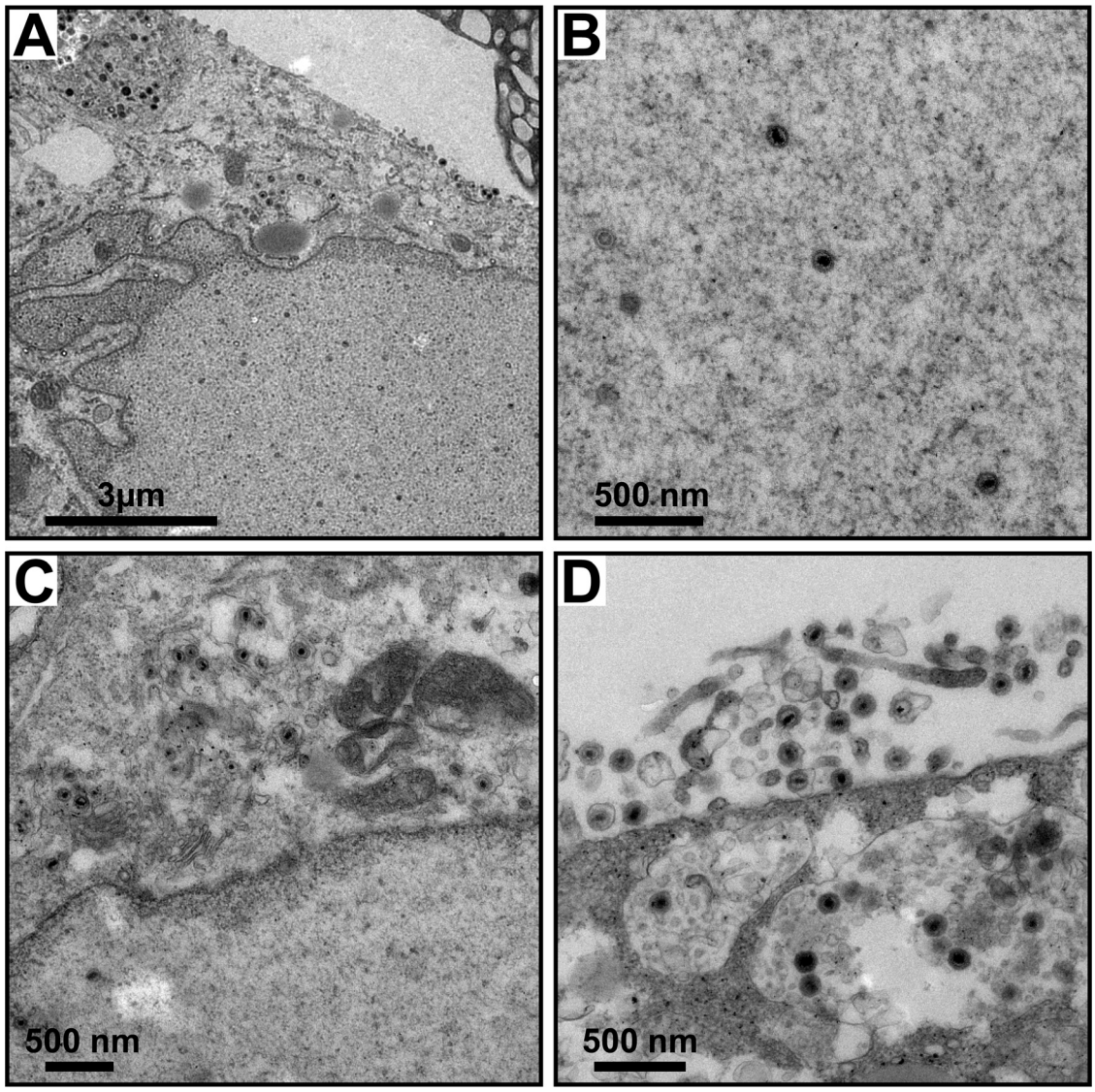

FIG 1 Morphology of TeHV-3. TH-1 cells were infected with strains 1976 and 4295. Six days postinfection, cells were processed for electron microscopy examination. In a blind test, it was not possible to differentiate the two strains. The images represent TH-1 cells infected with strain 1976. (A) General view of the various compartments of an infected cell; (B) nucleoplasm; (C) nucleus and cytoplasm; (D) cytoplasm and extracellular space.

Transmission electron microscopy. TH-1 cells were infected with TeHV-3 at an MOI of 0.2 PFU/cell and, at 6 days postinfection, processed for electron microscopic examination as described elsewhere (32).

Tortoises. Five-year-old Hermann's tortoises (*T. hermanni*) originating from a small colony bred in captivity were kept individually in terrariums (width by depth by height, 0.9 by 0.45 by 0.6 m). The environmental parameters were as follows. Relative humidity was maintained at 60 to 70%. Lighting was controlled automatically on standard 12-h light and 12-h dark circadian cycles, with a UVB light switched on during the 12-h light period. The temperature of an infrared basking spot was regulated at 29°C during the light period and 24°C during the dark period. A temperature gradient of approximately 6°C was present in the terrariums, with the basking spot being the warmest place. Fresh water and vegetables were provided daily. Clinical examinations of tortoises immediately prior to the experiments revealed that they were healthy. Experiments were preceded by an acclimatization period of 2 weeks.

Inoculation of tortoises with TeHV-3. Tortoises were sedated with alfaxolone (Dechra Veterinary Products) injected intravenously at a dose of 7 mg/kg of body weight. TeHV-3 was then inoculated by intranasal instillation of 1×10^5 PFU distributed equally between the two nostrils (total volume, 50 μl). The animals were examined twice daily until the end of the experiment. Animals expressing significant apathy, neurological signs, or respiratory distress were euthanized in accord with the endpoint defined by the local bioethics committee.

Ethics statement. The experiments, maintenance, and care of tortoises complied with the guidelines of the European Convention for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (CETS no. 123). The animal study was approved by the local ethics committee of the University of Liège, Belgium (laboratory accreditation no. LA 1610010, protocol no. 1217). All efforts were made to minimize suffering and to respect the 3Rs rule (reduction, refinement, and replacement).

Quantification of viral gene copies in organs by qPCR. DNA was isolated from 25 mg of organs stored at −80°C in RNAlater (Invitrogen) by using a DNA minikit (Qiagen). The viral genome was quantified by amplifying fragments of the TeHV-3 UL13 and *T. hermanni* β-actin

TABLE 1 Oligonucleotide primers

Description	Primer name	Sequence (5′–3′)	Coordinates[a]
For synthesis of probes for Southern blot analysis			
P1	TE2 F	ATACAGTCCGTGGGATCCAG	1303–1322
	TE2 R	CACGTGAGGCACATAGGAGA	1516–1535
P2	TE11 F	CAGAGGCTGAAGGAAACTGG	13285–13304
	TE11 R	TCCTCCCGCTATAGGAAACC	13477–13496
P3	TE12 F	AAGCCTGGTGGTACGATGAC	15747–15766
	TE12 R	GCAATCTCCGATAAGCTCCA	15960–15979
P4	UL1 F	TTTCCCGTACCTCTGTCTCG	120587–120606
	UL1 R	ATGAGATGTTGCTGCGACTG	120960–120979
P5	TE28 F	ATAGCGGCCGACAATGTAAC	134019–134038
	TE28 R	TGGCCCGAAGTATTTTTACG	134176–134195
P6	TE42 F	GTCTCGGTCCATGGCTATTC	151441–151460
	TE42 R	TCAGGGAGTAGTGGGTGGAG	151666–151685
For qPCR analysis (gene amplified)			
TeHV-3 UL13	UL13 F	CGCATCCGTCAGGAATCTAT	103686–103705
	UL13 R	GGTCCCTCGTCCAACTAACA	103772–103791
T. hermanni β-actin gene[b]	Beta-actin F	TCTGGTCGTACCACAGGTATTG	NA
	Beta-actin R	AGATCCAGACGGAGGATGG	NA
For analyzing specific features in the strain 1976 and strain 4295 genomes (targeted feature)[c]			
Reiteration at 6540–6842	TE5 F	AATACATGATACCAATCCCAGTTG	6473–6496
	TE5 R	CCAGAGGGGACACCGCAGATGACA	6902–6925
Reiteration at 27611–28227	TE19/1 F	ACTCCTGGCCACAGAACCAGTTGG	27559–27582
	TE19/1 R	TAAAAACATACCAGGAGGTTCCCA	28251–28274
Reiteration at 31883–32067	ORF13 F	ATCCAATACATATTTCACGG	31818–31837
	ORF13 R	CTGGTTCGAAACCTACGTATCGAG	32122–32145
Reiteration at 59762–59971	UL36 F	GTTGTTTACCCAGTTCCTGTACCA	59684–59707
	UL36 R	CATCTGGAAATAGTGTTGGCTATC	60040–60063
Reiteration at 97095–97300	TE22-TE23 F	CAAACGGCCATATCTCTTAG	97032–97051
	TE22-TE23 R	TAGAGCTCTGATCAATGTGTATAC	97387–97410
Reiteration at 121235–121565	UL1-IRS/1 F	TTTACAAAGCCGGGTGGAGCCTGG	121132–121155
	UL1-IRS/1 R	AGTCCCGGCATCGGCCGTGGTGGA	121629–121652
Duplication at 121745–121964 (ori)	UL1-IRS/2 F	TACAAGTACCTGATCGGGCT	121682–121701
	UL1-IRS/2 R	CCGAGATTTGGTACGGCTAGGACC	122396–122419
Reiteration at 126971–127795 and 151910–152308	TE25-RS1 F	ACGATTACGATGAGAGCACTGACA	126811–126834
	TE25-RS1 R	AGAACTGACCGGATTGGTGAACGA	127965–127988
Reiteration at 130962–131360 and 155475–156299	RS1-US F	GACCAGGCAGGTGCTTCATCCGTA	130866–130889
	RS1-US R	CTCCTAGCATTCCCATTGGC	131378–131397[d]
Reiteration at 131398–131576	RS1-TE26 F	GTCATGTCAACCAGCCAATG	131365–131384
	RS1-TE26 R	TTCATGAGGGTCACACTGAC	131607–131626
Reiteration at 150542–150717	TE42/1 F	TTCTGATATCCTGGGGACAT	150461–150480
	TE42/1 R	ATAGTAGTGTTGATTCGTCC	150754–150773
Reiteration at 151008–151141	TE42/2 F	AGAGGTCGCTGCTCTCTTAACTGA	150907–150930
	TE42/2 R	TAGAGGTAGAGGCTACGCCAGGTC	151370–151393
Duplication in RS1 (strain 4295)[d]	RS1 M F	ACGACATCTGCTATCTCAGTGCAC	129099–129122
	RS1 M R	TATACCAGGCTTTCGATGCGCTGG	130014–130038
Left genome terminus	Left terminus R	GCGAATCCGCAGGCAATCGCAACA	205–228
Left genome terminus[e]	Left terminus F	GCGATCCAAATACGCGTAGCGATT	122704–122727
Right genome terminus	Right terminus F	ACCCTCCGAAGAGCAGACCATGAG	160211–160234
For genotyping strain 4295 forms (targeted form or gene)			
M	TE4/1 F	GATGGGTATGGAACGTCACC	3790–3809
	TE16 R	CGGCCATGGTTAGAAAAAGA	22567–22586
m1	TE4/2 F	CAATCATCTGAGCGTTGGAA	4534–4553
	TE19/2 R	ATTCGTCCGTCACAGTAGGG	27514–27533
m2	TE3 F	AAAGTCCGCTCCTCTCATCA	2281–2300
	up TE12 R	GCCGCCTAATAGGTTCTTTG	15221–15240
TE8	TE8 F	CGAGCAGCCTAATTCAGACC	10046–10065
	TE8 R	AACGCCTTTCTGAACGAAGA	10236–10255

[a] Coordinates are listed in relation to the strain 1976 genome. NA, not applicable.
[b] Amplified fragment (primers omitted): 5′-TCT GGT CGT ACC ACA GGT ATT GTG ATG GAC TCT GGT GAT GGT GTC ACC CAC ACT GTG CCC ATC TAT GAA GGT TAT GCC CTC CCC CAC GCC ATC CTC CGT CTG GAT CT-3′.
[c] Principal primers only. Additional primers were used to generate confirmatory PCR products or sequence them. An adapter-specific primer was used in combination with virus-specific primers to locate the genome termini.
[d] Present in strain 4295 only. Coordinates correspond to the strain 1976 genome.
[e] U_L is inverted in a proportion of strain 4295 genomes.

genes, using real-time SYBR green-based PCR. The primers used are listed in Table 1. The quantitative real-time PCRs (qPCRs) were performed using a CFX96 Touch real-time PCR detection system and iTaq universal probe supermix as detection chemistry (Bio-Rad Laboratories). The master mix for qPCR consisted of 1× iTaq universal probe supermix, 200 nM each primer, and 200 ng sample DNA in a final volume of 15 μl. The UL13 amplification program included an initial denaturation step at 95°C for 3 min, followed by 50 cycles with a denaturation step at 95°C for 30 s, an annealing step at 58.5°C for 30 s, and an elongation step at 72°C for 30 s. The β-actin amplification program included an initial denaturation step at 95°C for 3 min, followed by 50 cycles with a denaturation step at 95°C for 30 s, an annealing step at 60°C for 30 s, and an elongation step at 72°C for 30 s. At the end of these amplification programs, the dissociation stage was performed (95°C for 10 s), and the melting curve was determined by increasing the temperature from 60 to 95°C at the rate of 0.1°C/s. All reactions were carried out in triplicate.

Data for validation of qPCR, i.e., efficiency (E), coefficient of determination (R^2), and slope, were analyzed by using Bio-Rad CFX Manager 3.1 software (Bio-Rad Laboratories) with the "auto" method (for UL13, E = 94.7%, $R^2 = 0.994$, slope = −3.456; for β-actin, E = 95.2%, $R^2 = 0.982$, slope = −3.443). The mean values of the number of TeHV-3 genome copies in various organs were compared by using the Kruskal-Wallis test. A P value of <0.05 was considered significant, and a P value of <0.01 was considered highly significant.

Histological analyses. Lung, spleen, brain (telencephalon), kidney, and liver from a mock-infected tortoise and from infected tortoises were dissected immediately after euthanasia, fixed in 10% buffered formalin, and embedded in paraffin blocks. Sections of 5 μm were stained with hematoxylin and eosin prior to microscopic analysis (34). For each sample, 10 randomly selected fields were examined in a blind test.

Phylogenetic analyses. Predicted amino acid sequences were obtained in this study or from GenBank. The raw phylogenetic data derived from them are available at http://dx.doi.org/10.17635/lancaster/researchdata/11. The sequences of herpesvirus DNA polymerases were aligned by using Muscle (35) in MEGA (36). A Bayesian tree was then constructed in BEAST (37), using the LG+Γ substitution model (38), a lognormal relaxed clock model (39), and a Yule speciation tree model (40). The tree was run to convergence, at which point all posterior probabilities on the nodes were 1. The amino acid sequence of TeHV-3 gene TE7, which encodes a semaphorin (SEMA)-related protein, was aligned with the sequences of SEMA-7s by using Muscle (35), before constructing a maximum likelihood tree in MEGA (36), using the LG+Γ substitution model (2). The amino acid sequence of TeHV-3 gene TE8, which encodes an interleukin-10 (IL-10)-related protein, was analyzed by using TreeAdder (41) with a previously published tree (42), in order to create a set of trees with the TeHV-3 TE8 sequence added at every possible position. The likelihood of each tree was assessed by using Tree-puzzle (43), and bootstrap significance was assessed by using Consel (44).

Structural analyses. Solved structures from PDB were selected for the SEMA and IL-10 protein families. Homology models for TeHV-3 SEMA and IL-10 were constructed in relation to these solved structures by using MOE 2014.09 (Chemical Computing Group, Montreal, Canada). The full data set for the homology modeling is available at http://dx.doi.org/10.17635/lancaster/researchdata/28. Briefly, an initial partial geometry was copied from the template chains in the solved structures, using all coordinates in which residue identity was conserved. Otherwise, only backbone coordinates were used. Based on this initial geometry, Boltzmann-weighted randomized modeling (45) was employed, with segment searching for regions that could not be mapped onto the initial geometry (46). A total of 100 models were constructed. On completion of segment addition, each model was energetically minimized in the AMBER99 force field (47). The highest-scoring intermediate model was then determined by the generalized Born/volume integral (GB/VI) methodology (48). The stereochemical quality of homology models was assessed using Ramachandran plots (49).

Nucleotide sequence accession numbers. The genome sequences of TeHV-3 strain 1976 and the three subclones (m1, m2, and M) of strain 4295 were deposited in NCBI GenBank under accession numbers KM924292, KR363629, KR363628, and KM924293, respectively.

RESULTS AND DISCUSSION

The initial goal of this study was to determine the sequence of the TeHV-3 genome. To achieve this, the two most-studied isolates (strains 1976 and 4295) were sequenced. Analysis of high-throughput sequence data for strain 1976 indicated the presence of a single population, whereas that for strain 4295 implied the presence of a mixture of three closely related populations, each being most simply interpreted as a deletion mutant derived from an ancestral genome similar to that of strain 1976.

Genome sequence of strain 1976. The filtered data set for strain 1976 consisted of 3,787,248 reads, of which 1,488,829 (39%) aligned with the finished sequence at an average coverage of 2,305 reads/nt. The genome is 160,358 bp in size and consists of a central, long unique region (U_L; 107,928 bp), extended at its right end by a short unique region (U_S; 20,375 bp) flanked by an inverted repeat (IR_S and TR_S; 8,536 bp) and at its left end by a third unique region (U_T; 12,595 bp) also flanked by an inverted repeat (TR_T and IR_T; 1,194 bp), yielding the overall configuration TR_T-U_T-IR_T-U_L-IR_S-U_S-TR_S (Fig. 2). Each 3′-end of the genome consists of a single, unpaired nucleotide complementary to the nucleotide at the other 3′-end. All alphaherpesvirus genome structures share the U_L-IR_S-U_S-TR_S component, in some cases with U_L flanked by an inverted repeat to produce TR_L-U_L-IR_L-IR_S-U_S-TR_S (2). However, the TR_T-U_T-IR_T component is novel, yielding a genome structure that has not been reported previously in the family *Herpesviridae*. The alphaherpesvirus genomes most reminiscent of TeHV-3 in this regard are equid herpesviruses 1 and 4 (genus *Varicellovirus*), which contain a small inverted repeat (87 and 86 bp, respectively) at the left end of U_L, separated from its counterpart by a short sequence (944 and 667 bp, respectively) (50, 51). However, the sizes of TR_T/IR_T and U_T in TeHV-3 are much greater than in these viruses, and each contains protein-coding sequences (see below). The average nucleotide composition of the strain 1976 genome is 46% G+C, with TR_T/IR_T, U_T, U_L, TR_S/IR_S, and U_S being 65, 45, 44, 55, and 45%, respectively. Further details of the strain 1976 sequence, including an annotation of predicted protein-coding content (see below), are available in GenBank. Four partial sequences from strain 1976 have been published previously, and all are identical to the corresponding sections of the genome sequence. These include 8,667 bp extending from within UL40 to within UL36 (GenBank accession AY338245) and much shorter sequences from UL39 (DQ343900), UL5 (DQ343892), and UL30 (DQ343881) (6).

In alphaherpesvirus genomes, unique regions flanked by inverted repeats are typically present in either orientation in virion DNA, giving rise to the presence of more than one genome isomer (2). Consequently, the orientations of U_T and U_S in the strain 1976 genome were investigated by Southern blotting (Fig. 3). Figure 3A illustrates the four possible arrangements of U_T-U_L and U_L-U_S and the restriction endonuclease fragments potentially generated by digestion with EcoRI or KpnI. Hybridizations performed with appropriate probes showed that U_T and U_S are present in either orientation, whereas U_L, which is not flanked by an inverted repeat, is present in a single orientation (Fig. 3B). These results indi-

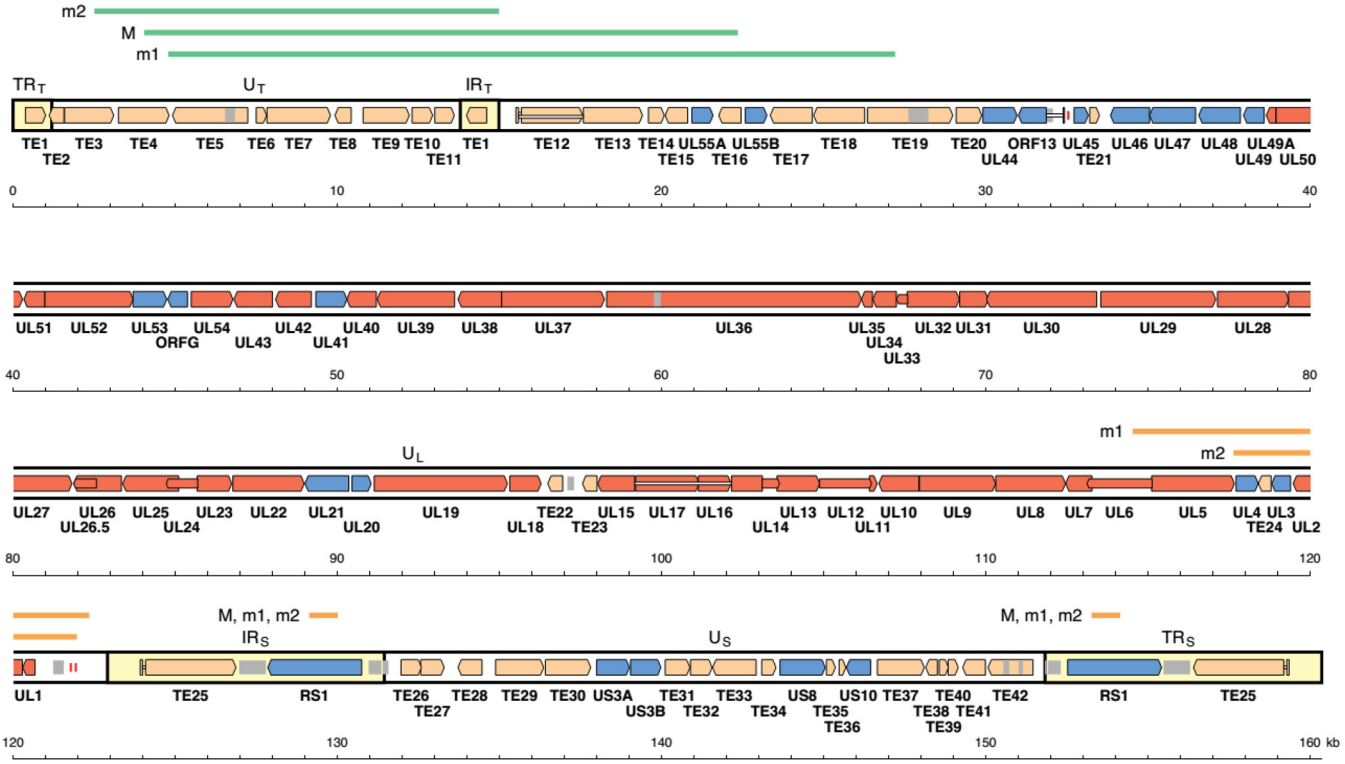

FIG 2 Map of the TeHV-3 strain 1976 genome. The unique regions (U_T, U_L, and U_S) are in white, and the inverted repeats (TR_T, IR_T, IR_S, and TR_S) are in light yellow. Predicted functional ORFs are depicted by colored arrows, with nomenclature below them. Red shading indicates ORFs inherited from the ancestor of the alpha-, beta-, and gammaherpesviruses. Blue shading indicates ORFs that have orthologs in mammalian or avian alphaherpesviruses. Light orange shading indicates the remaining ORFs. Introns are shown as narrow white bars. Reiterations are shown by gray shading within or between ORFs, and three copies of ori are indicated by vertical red bars. The deletions and the duplications present in the m1, m2, and M forms of TeHV-3 strain 4295 are marked above the genome by horizontal green and orange bars, respectively.

cate that TeHV-3 virion populations contain a mixture of four genome isomers differing in the relative orientations of U_T and U_S.

Genome sequence of strain 4295. The filtered data set for strain 4295 consisted of 9,772,400 reads, of which 3,106,488 (32%) aligned with the finished sequence at an average coverage of 5,195 reads/nt. Analysis of the data indicated that the DNA contained a mixture of three related genome populations, each most simply interpreted as being a deleted form of a genome similar to that of strain 1976. The extents of the deletions in relation to strain 1976 are marked by horizontal green bars in Fig. 2. The three genome forms were named m1 and m2 (the minor forms, estimated from counting reads representing the novel junction, representing approximately 23 and 11% of the population, respectively) and M (the major form representing approximately 66% of the population). This interpretation was confirmed by subcloning the m1, m2, and M forms by limiting dilution assays and sequencing their genomes at average coverages of 25, 46, and 22 reads/nt, respectively. Compared with the strain 1976 sequence, the m1, m2, and M forms exhibit large, partially overlapping deletions of 22,424, 12,485, and 18,315 bp, respectively, extending from U_T (in the orientation shown in Fig. 2), across IR_T, to U_L. No reads matching the region in the strain 1976 genome that corresponds to the region absent from all three forms (nt 4780 to 14989 in strain 1976) were detected in strain 4295 prior to subcloning, indicating that none of the putative parental genome remained. In addition, the data supported the presence of two sizeable duplications in strain 4295 prior to subcloning and in the subclones, one located near the right end of U_L (present in the m1 and m2 forms) and the other in IR_S and TR_S (present in all three forms). The extents of these duplications relative to strain 1976 are marked by horizontal orange bars in Fig. 2. Further details of the sequences of the strain 4295 subclones, including annotations of coding content, are available in GenBank. No reads diagnostic of the deletions and duplications in strain 4295 (i.e., reads representing the novel junctions) were detected in strain 1976. The genomes of strains 1976 and 4295 (the M form) are closely related, differing by 193 substitutions (including IR_S but not TR_S and not counting insertions or deletions).

General features of the TeHV-3 genome. Standard bioinformatic approaches were taken to predict the locations of open reading frames (ORFs) in the strain 1976 genome that encode functional proteins (for an example, see reference 30). In general, ORFs were included that potentially encode proteins of 100 or more amino acid residues, that do not extensively overlap ORFs predicted to be functional (i.e., ORFs encoding proteins that are similar to proteins of known function or have features suggesting function, such as hydrophobic domains), and that are located appropriately in relation to potential mRNA polyadenylation signals. Particular attention was paid to ORFs that have counterparts in other herpesviruses. Four ORFs potentially encoding proteins of fewer than 100 residues were also added because they are related to recognized herpesvirus proteins (UL11) or have other distin-

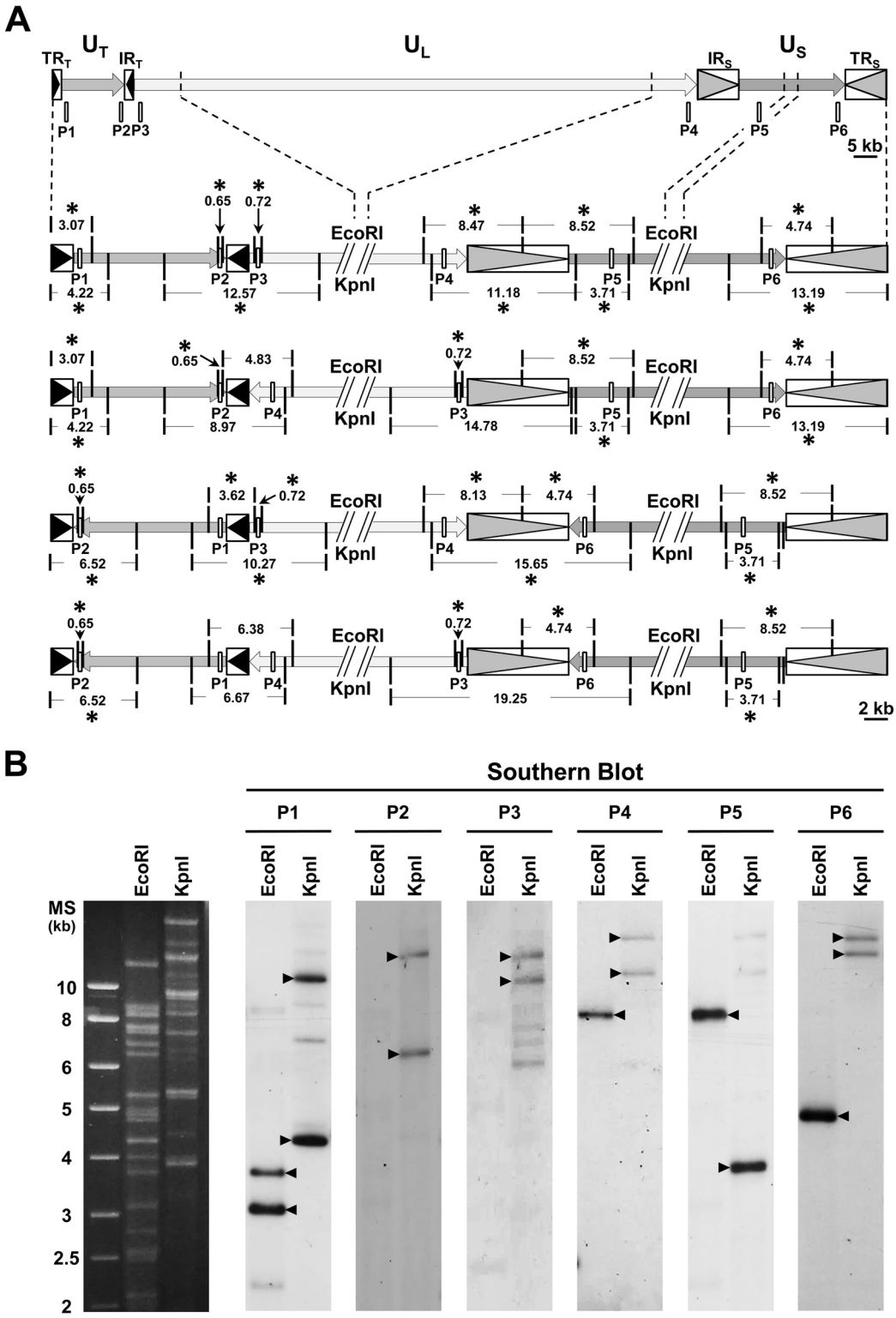

FIG 3 Relative orientations of unique regions in the TeHV-3 strain 1976 genome. (A) A schematic representation of the genome is shown at the top, the orientations of U_T, U_L, and U_S corresponding to those in Fig. 2. Below this, the four possible combinations of the orientations of U_T-U_L and U_L-U_S are presented, with the majority of U_L omitted. For each combination, the sizes (in kilobases) of restriction endonuclease fragments at or near the genome ends or the U_T-U_L and U_L-U_S junctions are shown (EcoRI above the genome and KpnI below). White bars (P1 to P6) represent the positions of the probes used for hybridization. Asterisks highlight restriction endonuclease fragments for which a positive band was observed in Southern blot analyses. (B) The panels on the right show a Southern blot analysis of TeHV-3 strain 1976 DNA digested with EcoRI or KpnI and hybridized to probes P1 to P6. Black arrowheads indicate all bands compatible with the predicted fragments. The panel on the left shows an ethidium bromide-stained profile of marker fragments (MS) and strain 1976 DNA digested with EcoRI or KpnI. The 0.65- and 0.72-kb fragments were detected but are not visible on these images, which have been restricted to fragments above 2 kb.

guishable features (TE35, TE36, and TE39). Splicing was predicted in five ORFs (TE12, TE13, ORF13, UL15, and TE25). The first ATG in each ORF was assigned as the initiation codon, except in cases in which use of a subsequent ATG was supported by alignments with related proteins or in which it provided a putative signal peptide.

The analysis indicated that the strain 1976 genome encodes a total of 107 predicted functional ORFs (Fig. 2 and Table 2), three of which are duplicated in the inverted repeats (TE1 in TR_T/IR_T, and TE25 and RS1 in TR_S/IR_S). These 107 ORFs are conserved in strain 4295, except for those affected by the deletions, and TE35, which is frameshifted (and therefore considered marginal in strain 1976). A total of 19 ORFs belong to six families of paralogous genes: the TE3 family with nine members, and the TE15, TE22, TE27, UL55, and US3 families, each with two members. The genome also contains three copies of a potential origin of DNA replication (ori), identified as an A+T-rich region capable of forming a hairpin structure and containing characteristic sequence motifs involved in binding the UL9 DNA replication origin-binding helicase (52–54) (Fig. 2). Approximately 37% of strain 1976 genomes appear to contain a 154-bp deletion that results in the absence of one of the copies of ori from the region between UL1 and TE25. This deletion also appears to be present in approximately 50% of strain 4295 genomes prior to subcloning, but their distribution in this strain and in subclones m1 and m2 was not determined with certainty because of the ambiguity caused by duplications in this region.

Relationships between TeHV-3 and other herpesviruses. The strain 1976 ORFs are shown in three categories in Fig. 2, according to their conservation in other herpesviruses. The first category (shaded red) consists of the 44 genes that are thought to have been inherited from the ancestor of the alpha-, beta-, and gammaherpesviruses. The second category (shaded blue) comprises the 21 additional ORFs (not counting duplicates) that have orthologs in alphaherpesviruses of mammals or birds. The third category (shaded orange) contains the remaining 42 ORFs (TE1 to TE42), all of which lack counterparts in other herpesviruses, except for TE31, which has an ortholog in the turtle alphaherpesvirus ChHV-5 (4). This category of TeHV-3-specific gene includes the TE3, TE15, TE22, and TE27 families, as well as TE7, TE8, and TE11 genes, which are similar to SEMAs, IL-10s, and C-type lectin genes, respectively.

The full list of TeHV-3 strain 1976 ORFs is provided in Table 2, including those having counterparts in ChHV-5 (marked with asterisks in the first column). The order of conserved ORFs in the TeHV-3 genome is the same as that in ChHV-5, except for the absence or disruption of 14 ORFs in the latter (UL44-UL51, UL54, UL39, UL40, and UL13) and the inversion of UL55 (of which TeHV-3 has two counterparts, UL55A and UL55B). Comparison with other alphaherpesvirus lineages (in the mammalian alphaherpesvirus genera *Simplexvirus* and *Varicellovirus* and the avian alphaherpesvirus genera *Mardivirus* and *Iltovirus*) (55) indicates that the ancestral alphaherpesvirus contained conserved ORFs arranged in the same order as in TeHV-3 from UL43 rightwards. The ancestral state from UL54 leftwards is more difficult to discern, as this region appears to have undergone rearrangements in various lineages.

Phylogenetic analysis of short regions of the genome has indicated previously that TeHV-3 is an alphaherpesvirus (6). This conclusion is supported from the complete genome sequence by the presence in TeHV-3 of 21 ORFs that have orthologs in alphaherpesviruses but not in beta- or gammaherpesviruses (Fig. 2) and by phylogenetic analysis of the DNA polymerase (Fig. 4A). This analysis also indicates that the closest known relative of TeHV-3 is ChHV-5, the current sole member of the genus *Scutavirus* (1). However, despite their relationship, the two viruses do not share the same genome structure, ChHV-5 apparently adopting the simpler U_L-IR_S-U_S-TR_S arrangement (4). Given the relatively large phylogenetic distance between TeHV-3 and ChHV-5, it is a matter of judgment whether TeHV-3 should be classified as a new species into the same genus as ChHV-5 or into a new genus. Since these are the only relevant viruses that have been examined in sufficient detail and both infect Testudines, we recommend the former as the safer option. Regardless of the eventual taxonomical outcome, the results establish a robust phylogenetic relationship between tortoise and turtle herpesviruses.

Phylogenic analysis and homology modeling of the TeHV-3 TE7 and TE8 proteins. The finding that the TeHV-3 TE7 and TE8 genes encode SEMA and IL-10 homologs, respectively, is particularly interesting for the following reasons. First, SEMA homologs have been reported in the family *Herpesviridae* only in the *Macavirus* genus of subfamily *Gammaherpesvirinae* (56). Second, although numerous IL-10 homologs have been described in the subfamilies *Betaherpesvirinae* and *Gammaherpesvirinae* (42), they have not been reported in the subfamily *Alphaherpesvirinae*. We performed phylogenetic analyses to determine whether the origins of the TeHV-3 SEMA and IL-10 homologs could be traced. We also performed homology modeling analyses, in order to assess whether these viral genes encode functional homologs of the cellular genes.

A maximum likelihood phylogenetic tree placed the TeHV-3 SEMA (the TE7 protein) as the nearest neighbor to a cluster of poxvirus SEMAs, with low bootstrap confidence (Fig. 4B). No viral SEMA is close enough to any cellular counterpart to justify the deduction of an evolutionary scenario involving a recent gene transfer from host to virus. In contrast, as previously described (42), some viral IL-10s (in human herpesvirus 4 [HHV-4], equine herpesvirus 2 [EHV-2], and Orf virus [ORFV]) cluster phylogenetically with cellular IL-10s, implying relatively recent gene transfers and a degree of functional conservation (Fig. 4C). However, other viral IL-10s (including that of TeHV-3) are more divergent from host IL-10s, implying earlier gene transfer events with subsequent genetic drift and possible functional divergence. As a consequence, no inference concerning the evolutionary history of the TeHV-3 IL-10 (the TE8 protein) can be made, other than the speculation that it originated from a host in a relatively ancient transfer event, as appears to be the case for many other viral IL-10s (42).

We investigated the possibility of functional divergence by using homology modeling of the TE7 and TE8 proteins in comparison with cellular SEMAs and IL-10s, respectively (Fig. 5 and 6). The templates were selected for superposition on the basis of the lowest root mean square deviation (RMSD) values exhibited by the top BLAST hits. This process indicated that mouse SEMA-3A and human IL-10 were the best templates (Table 3). It should not be inferred from these choices that we believe the TE7 protein to be an ortholog of mouse SEMA-3A or the TE8 protein an ortholog of human IL-10. Indeed, the top BLAST hit of the TE7 protein among the cellular SEMAs was SEMA-7A; hence the use of this group of proteins in Fig. 4B. On the basis of the templates selected,

TABLE 2 Features of predicted functional protein-coding regions in the TeHV-3 strain 1976 genome

Gene name[a]	Gene family	Protein size (no. of residues)	Protein name[b]	Protein feature(s)
TE1		207	Protein TE1	
TE2		152	Protein TE2	
TE3	TE3	505	Protein TE3	Contains signal peptide
TE4	TE3	520	Protein TE4	Contains signal peptide
TE5	TE3	774	Protein TE5	Contains signal peptide
TE6		106	Protein TE6	Contains potential transmembrane domain
TE7		653	Semaphorin	Type 1 membrane protein
TE8		168	Interleukin-10	Contains signal peptide
TE9	TE3	473	Membrane protein TE9	Type 1 membrane protein
TE10		210	Protein TE10	
TE11		200	Protein TE11	Contains potential transmembrane domain; similar to C-type lectins
TE1		207	Protein TE1	
TE12	TE3	658	Protein TE12	Contains signal peptide
TE13	TE3	636	Protein TE13	Contains signal peptide
TE14		162	Protein TE14	
TE15	TE15	231	Protein TE15	
UL55A	UL55	218	Nuclear protein UL55A	
TE16	TE15	229	Protein TE16	
UL55B*	UL55	221	Nuclear protein UL55B	
TE17		432	Protein TE17	
TE18		525	Protein TE18	
TE19		876	Membrane protein TE19	Type 1 membrane protein; contains immunoglobulin domains
TE20		262	Membrane protein TE20	Type 1 membrane protein; contains immunoglobulin domain
UL44		350	Envelope glycoprotein C	Type 1 membrane protein; binds cell surface heparan sulfate; binds complement C3b to block neutralization; involved in cell attachment
ORF13		302	Thymidylate synthase	Involved in nucleotide metabolism
UL45		146	Membrane protein UL45	Type 2 membrane protein; tegument-associated; possibly involved in membrane fusion
TE21		100	Protein TE21	Contains potential transmembrane domain
UL46		403	Tegument protein VP11/12	Modulates transactivating tegument protein VP16; possibly involved in gene regulation
UL47		473	Tegument protein VP13/14	Modulates transactivating tegument protein VP16; RNA-binding protein; possibly involved in gene regulation
UL48		428	Transactivating tegument protein VP16	Transactivates immediate early genes; involved in gene regulation; involved in virion morphogenesis
UL49		212	Tegument protein VP22	Involved in virion morphogenesis; possibly involved in RNA transport to uninfected cells
UL49A		100	Envelope glycoprotein N	Type 1 membrane protein; complexed with envelope glycoprotein M; involved in virion morphogenesis; involved in membrane fusion
UL50		447	Deoxyuridine triphosphatase	Involved in nucleotide metabolism
UL51		216	Tegument protein UL51	Involved in virion morphogenesis
UL52*		906	Helicase-primase primase subunit	Involved in DNA replication
UL53*		346	Envelope glycoprotein K	Type 3 membrane protein; 4 transmembrane domains; involved in virion morphogenesis; involved in membrane fusion
ORFG*		201	Protein IG	Contains a possible zinc-binding domain
UL54		953	Multifunctional expression regulator	RNA-binding protein; shuttles between nucleus and cytoplasm; inhibits pre-mRNA splicing; exports virus mRNA from nucleus; exerts most effects post-transcriptionally; involved in gene regulation; involved in RNA metabolism and transport
UL43*		396	Envelope protein UL43	Type 3 membrane protein; possibly involved in membrane fusion
UL42*		368	DNA polymerase processivity subunit	dsDNA-binding protein; involved in DNA replication
UL41*		319	Tegument host shutoff protein	mRNA-specific RNase; involved in cellular mRNA degradation
UL40		304	Ribonucleotide reductase subunit 2	Involved in nucleotide metabolism
UL39		796	Ribonucleotide reductase subunit 1	Involved in nucleotide metabolism
UL38*		447	Capsid triplex subunit 1	Complexed 1:2 with capsid triplex subunit 2 to connect capsid hexons and pentons; involved in capsid morphogenesis
UL37*		1056	Tegument protein UL37	Complexed with large tegument protein; involved in virion morphogenesis

(Continued on following page)

TABLE 2 (Continued)

Gene name[a]	Gene family	Protein size (no. of residues)	Protein name[b]	Protein feature(s)
UL36*		2619	Large tegument protein	Complexed with tegument protein UL37; ubiquitin-specific protease (N-terminal region); involved in capsid transport
UL35*		114	Small capsid protein	Located externally on capsid hexons; involved in capsid morphogenesis; possibly involved in capsid transport
UL34*		242	Nuclear egress membrane protein	Type 2 membrane protein; interacts with nuclear egress lamina protein; involved in nuclear egress
UL33*		115	DNA-packaging protein UL33	Interacts with DNA packaging terminase subunit 2; involved in DNA encapsidation
UL32*		537	DNA-packaging protein UL32	Involved in DNA encapsidation; possibly involved in capsid transport
UL31*		295	Nuclear egress lamina protein	Interacts with nuclear egress membrane protein; involved in nuclear egress
UL30*		1132	DNA polymerase catalytic subunit	Involved in DNA replication
UL29*		1183	Single-stranded DNA-binding protein	Contains a zinc finger; involved in DNA replication; possibly involved in gene regulation
UL28*		730	DNA-packaging terminase subunit 2	Involved in DNA encapsidation
UL27*		827	Envelope glycoprotein B	Type 1 membrane protein; possible membrane fusogen; binds cell surface heparan sulfate; involved in cell entry; involved in cell-to-cell spread
UL26.5*		233	Capsid scaffold protein	Clipped near C terminus; involved in capsid morphogenesis
UL26*		490	Capsid maturation protease	Serine protease (N-terminal region); minor scaffold protein (remainder of protein, clipped near C terminus); involved in capsid morphogenesis
UL25*		575	DNA-packaging tegument protein UL25	Located on capsid near vertices; possibly stabilizes the capsid and retains the genome; involved in DNA encapsidation
UL24*		324	Nuclear protein UL24	
UL23*		355	Thymidine kinase	Involved in nucleotide metabolism
UL22*		733	Envelope glycoprotein H	Type 1 membrane protein; possible membrane fusogen; complexed with envelope glycoprotein L; involved in cell entry; involved in cell-to-cell spread
UL21*		450	Tegument protein UL21	Interacts with microtubules; involved in virion morphogenesis
UL20*		196	Envelope protein UL20	Type 3 membrane protein; 4 transmembrane domains; involved in virion morphogenesis; involved in membrane fusion
UL19*		1369	Major capsid protein	6 copies form hexons, 5 copies form pentons; involved in capsid morphogenesis
UL18*		320	Capsid triplex subunit 2	Complexed 2:1 with capsid triplex subunit 1 to connect capsid hexons and pentons; involved in capsid morphogenesis
TE22	TE22	153	Protein TE22	
TE23	TE22	150	Protein TE23	
UL15*		697	DNA-packaging terminase subunit 1	Contains an ATPase domain; involved in DNA encapsidation
UL17*		652	DNA-packaging tegument protein UL17	Capsid-associated; involved in DNA encapsidation; involved in capsid transport
UL16*		335	Tegument protein UL16	Possibly involved in virion morphogenesis
UL14*		177	Tegument protein UL14	Involved in virion morphogenesis
UL13		437	Tegument serine/threonine protein kinase	Involved in protein phosphorylation
UL12*		533	DNase	Involved in DNA processing
UL11*		79	Myristylated tegument protein	Envelope associated; involved in virion morphogenesis
UL10*		412	Envelope glycoprotein M	Type 3 membrane protein; 8 transmembrane domains; complexed with envelope glycoprotein N; involved in virion morphogenesis; involved in membrane fusion
UL9*		784	DNA replication origin-binding helicase	Involved in DNA replication
UL8*		716	Helicase-primase subunit	Involved in DNA replication
UL7*		271	Tegument protein UL7	Involved in virion morphogenesis
UL6*		669	Capsid portal protein	Dodecamer located at one capsid vertex in place of a penton; involved in DNA encapsidation
UL5*		851	Helicase-primase helicase subunit	Involved in DNA replication
UL4*		227	Nuclear protein UL4	Colocalizes with regulatory protein ICP22 and nuclear protein UL3 in small, dense nuclear bodies
TE24		131	Protein TE24	
UL3*		193	Nuclear protein UL3	Colocalizes with regulatory protein ICP22 and nuclear protein UL4 in small, dense nuclear bodies

(Continued on following page)

TABLE 2 (Continued)

Gene name[a]	Gene family	Protein size (no. of residues)	Protein name[b]	Protein feature(s)
UL2*		270	Uracil-DNA glycosylase	Involved in DNA repair
UL1*		127	Envelope glycoprotein L	Contains signal peptide; complexed with envelope glycoprotein H; involved in cell entry; involved in cell-to-cell spread
TE25	TE3	953	Protein TE25	Contains signal peptide
RS1*		965	Transcriptional regulator ICP4	Involved in gene regulation
TE26		210	Protein TE26	
TE27	TE27	241	Protein TE27	Contains potential transmembrane domain
TE28		246	Protein TE28	Contains potential transmembrane domain
TE29	TE3	494	Protein TE29	Contains signal peptide
TE30	TE3	465	Protein TE30	Contains signal peptide
US3A*	US3	338	Serine/threonine protein kinase US3A	
US3B*	US3	312	Serine/threonine protein kinase US3B	
TE31*		253	Membrane protein TE31	Type 1 membrane protein
TE32		222	Membrane protein TE32	Type 1 membrane protein; contains Ig domain
TE33		448	Protein TE33	Contains signal peptide
TE34		147	Protein TE34	
US8*		465	Envelope glycoprotein E	Type 1 membrane protein; involved in cell-to-cell spread
TE35		92	Protein TE35	Contains signal peptide
TE36		69	Protein TE36	Contains potential transmembrane domain
US10*		248	Virion protein US10	
TE37		481	Protein TE37	Contains signal peptide
TE38		120	Protein TE38	
TE39		94	Protein TE39	Contains potential transmembrane domains
TE40		103	Protein TE40	
TE41	TE27	233	Protein TE41	
TE42		462	Protein TE42	
RS1*		965	Transcriptional regulator ICP4	Involved in gene regulation
TE25	TE3	953	Protein TE25	Contains signal peptide

[a] Genes are listed as they are ordered in the genome. Duplicates (TE1, RS1, and TE25) are included. Asterisks mark genes that have orthologs in ChHV-5. Information on conservation in other herpesviruses is available in Fig. 2.

[b] The names of conserved proteins and functional annotations are derived from NCBI reference sequences (http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?taxid=10292).

modeling clearly demonstrates that the TE7 protein (Fig. 5B) and the TE8 protein (Fig. 6B) are both capable of assuming backbone structures highly similar to their cellular counterparts, despite having a high degree of sequence divergence. Both TE7 and TE8 homology models contain some residues that are suboptimal in terms of their stereochemistry (Fig. 5C and 6C). However, these are mostly at turns in the backbone or loop regions (Fig. 5B and 6B). Major secondary structural regions, by contrast, are well formed in terms of their stereochemistry. We conclude that functional divergence is unlikely to have occurred to an extent such that these two viral proteins no longer operate in ways analogous to their cellular equivalents, and therefore that the TE7 and TE8 proteins are likely to have maintained SEMA and IL-10 functions. This conclusion is further supported by key residues of TE7 and TE8 proteins. The three-dimensional structure of SEMA-7A bound to its receptor PLEXIN-C1 has been resolved, and key residues in the SEMA sequence for the interaction with its receptor have been identified (57). Interestingly, most of these residues present in the blade 3, 4c-4d loop, and extrusion helix 2 regions were conserved in TE7. Among the residues conserved in TeHV-3, only one of them (Ile238) had a suboptimal stereochemistry (Fig. 5B). Together, these data suggested that the TeHV-3 TE7 protein might also signal through Plexin-C1 and function similarly to other viral SEMAs (56). As for the TE8 protein, its sequence exhibits most of the residues found in the two family signature motifs characteristic of all cellular IL-10s: L-[FILMV]-X3-[ILV]-X3-[FILMV]-X5-C-X5-[ILMV]-[ILMV]-X(3)-L-X2-[IV]-[FILMV] and KA-X2-E-X-D-[ILV]-[FLY]-[FILMV]-X2-[ILMV]-[EKQZ] (residues that are not conserved in the TE8 protein are underlined).

Requirement *in vitro* of the genome regions deleted in the three forms of strain 4295. The m1, m2, and M forms of strain 4295 exhibit partially overlapping deletions in their genomes (Fig. 2). The observation that all three forms lack the region corresponding to coordinates 4780 to 14989 in strain 1976 indicates that this 10,210-bp sequence is not essential for viral replication *in vitro*. As the three forms resulted from a coculture, it is possible that each of them requires complementation by the others in order to provide the functions that have been lost. To test this hypothesis, strain 4295 was subcloned by limiting dilution, and the subclones were analyzed by PCR (Fig. 7A and B). This experiment demonstrated that the three forms are capable of growing *in vitro*, despite relatively large deletions. To determine whether the deletions have quantitative effects on viral growth, the three subclones were compared using a multistep growth assay. They replicated comparably with each other and strain 1976, thereby demonstrating that none of the deleted genes affected viral replication in the assay used (Fig. 7C). Finally, the morphogenic properties of strains 1976 and 4295 were compared by electron microscopic examination of infected TH-1 cells (Fig. 1). In a blind test, it was

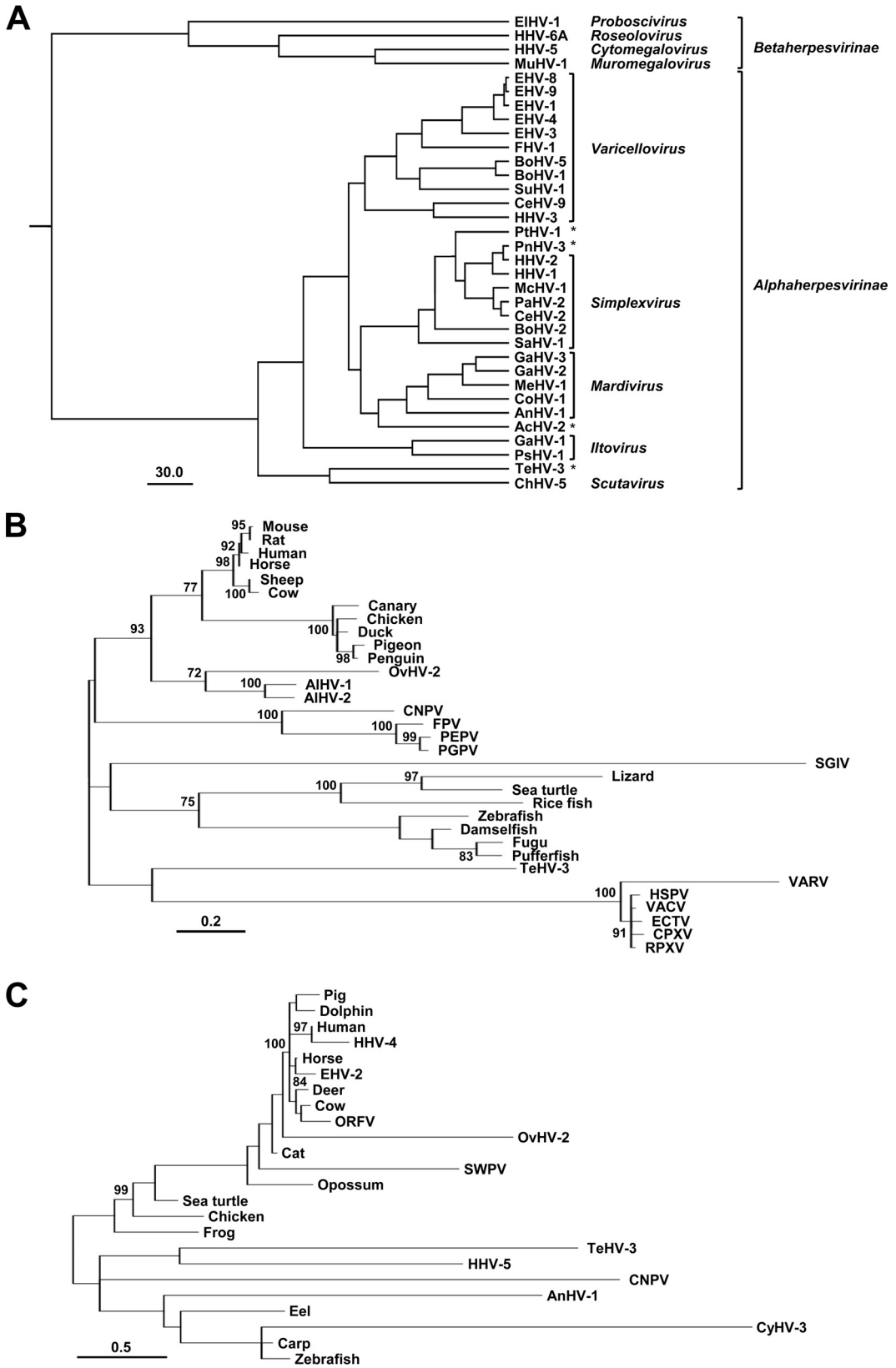

FIG 4 Phylogenetic analyses. In each panel, the scale indicates substitutions per site. Abbreviations for herpesvirus names: El, elephantid; H, human; Mu, murid; E, equid; F, felid; Bo, bovid; S, suid; Ce, cercopithecine; Pt, pteropodid; Pn, panine; Mc, macacine; Pa, papiine; Sa, saimiriine; Ga, gallid; Me, meleagrid; Co, columbid; An, anatid; Ac, accipitrid; Ps, psittacid; Te, testudinid; Ch, chelonid; Ov, ovine; Al, alcelaphine; Cy, cyprinid; and HV, herpesvirus (followed by a hyphenated number). Other abbreviations: CNPV, canarypox virus; FPV, fowlpox virus; PEPV, penguin poxvirus; PNPV, pigeon poxvirus; SGIV, Singapore grouper iridovirus; VARV, variola virus; HSPV, horse poxvirus; VACV, vaccinia virus; ECTV, ectromelia virus; CPXV, cowpox virus; RPXV, rabbit poxvirus; ORFV, ORF virus; and SWPV, swinepox virus. (A) Bayesian tree for herpesvirus DNA polymerase proteins. All nodes have posterior probabilities of 1. Viruses that have not yet been classified are marked by asterisks (1). (B) Maximum likelihood tree for vertebrate SEMA-7A proteins and their viral homologs. Node bootstrap values greater than 70 are marked. (C) Maximum likelihood tree for IL-10 homologs. Node bootstrap values greater than 70 are marked.

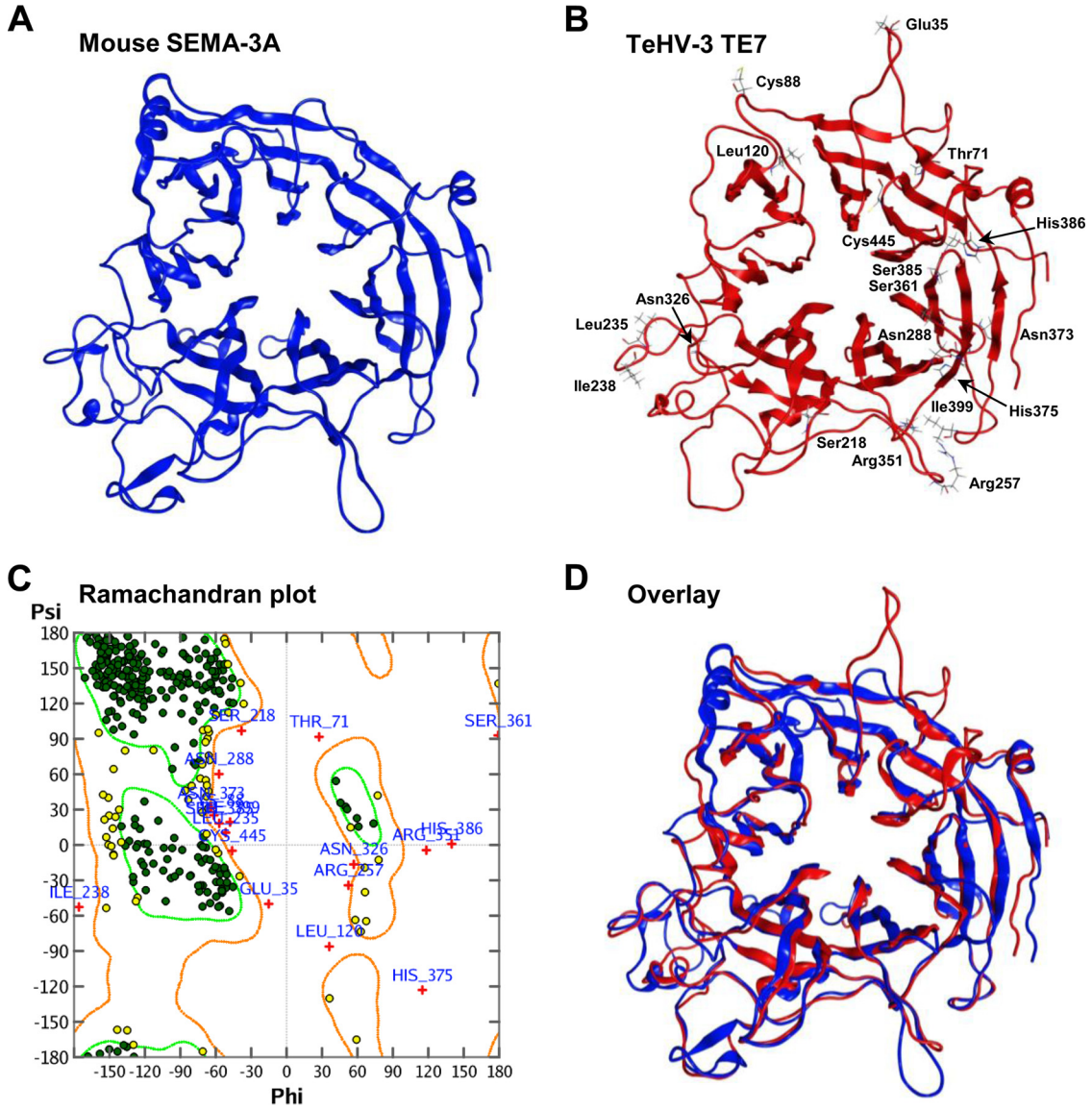

FIG 5 Homology model of the TeHV-3 SEMA (TE7) protein. (A) Mouse SEMA-3A PDB 1Q47. (B) Homology model of TeHV-3 TE7 constructed from PDB 1Q47 with suboptimal residues shown. (C) Ramachandran plot of homology model used to identify suboptimal residues. (D) Mouse SEMA-3A PDB 1Q47 (blue) superposed with the homology model of TeHV-3 TE7 (red).

not possible to differentiate the two strains, both of which exhibited all the features typical of TeHV-3 and herpesviruses in general.

Pathogenesis of strain 4295. The tortoises that are most sensitive to TeHV-3 infection belong to the genus *Testudo* (58). All species in this genus are protected by the Convention on International Trade in Endangered Species of wild fauna and flora (CITES). Consequently, their use in *in vivo* experiments is highly restricted and, moreover, carries the mandatory condition that the scientific objective of any such experiments must contribute to species conservation. As indicated above, no safe and efficacious vaccine is yet available against TeHV-3. The observation that strain 4295 consists of a mixture of three deletion mutants prompted us to test the potential of this mixture as an attenuated vaccine. With this goal in mind, three tortoises were inoculated with strain 4295 (the same passage as the one used for sequencing and which was shown to be a mixture of the m1, m2, and M forms) by intranasal instillation, and one sentinel tortoise was mock-infected (Fig. 8). Our intention was to observe the animals for 2 months, in order to assess the safety of the inoculated material and then to challenge them with strain 1976 and evaluate the immune protection conferred. However, all three infected tortoises became apathetic at about 20 days postinoculation (dpi) and then reduced feeding progressively, to become anorexic between 22 and 29 dpi. Significant nasal discharge was observed by 30 to 37 dpi, with mild blepharoedema. Diphtheritic plaques were observed in the buccal cavity of one tortoise (4295/41 D). The animals were euthanized at 38 and 41 dpi due to extreme weakness (defined as one of the endpoints by the local ethics committee). The mock-infected tortoise did not show any symptoms throughout the course of the experiment. It was euthanized at 75 dpi in order to serve as a negative control for further analyses (see below).

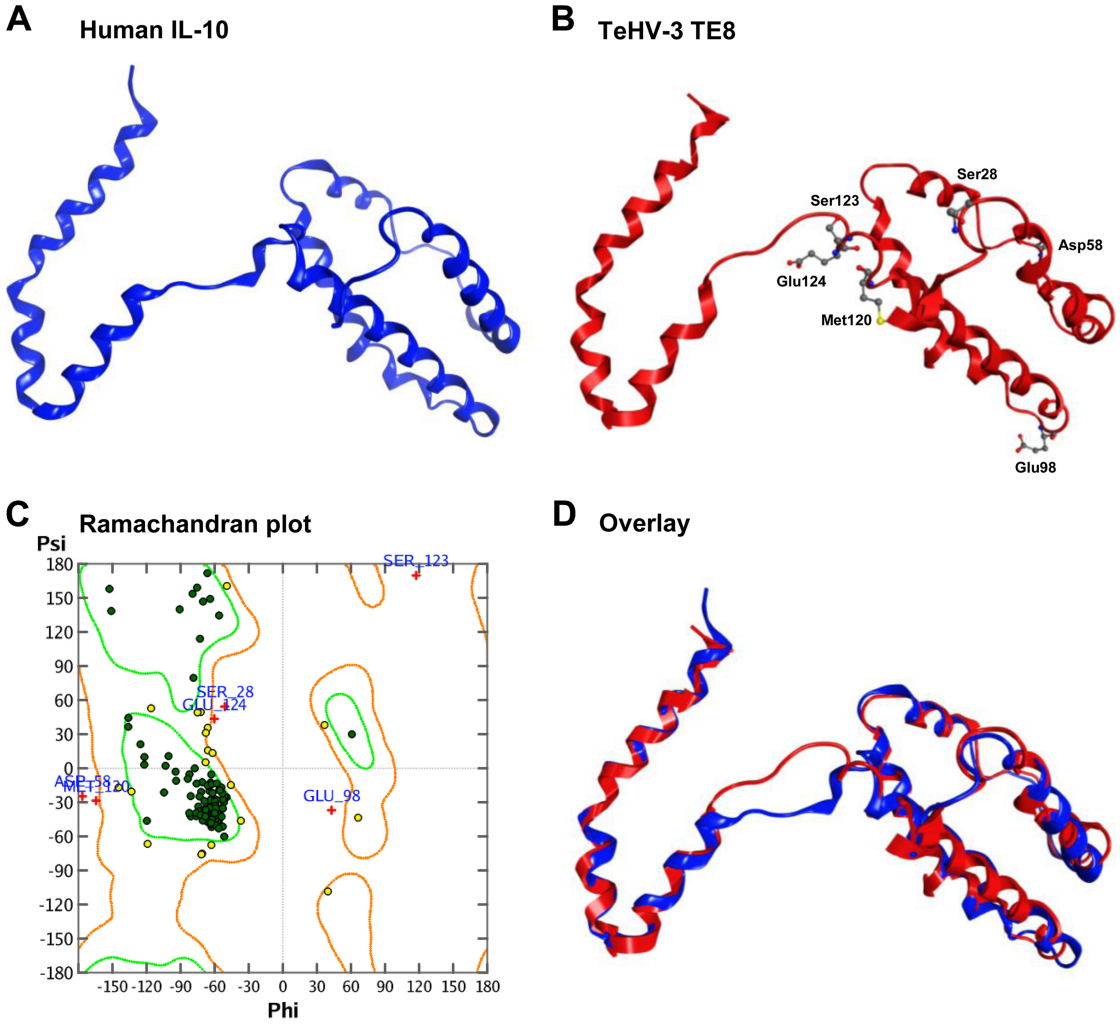

FIG 6 Homology model of the TeHV-3 IL-10 (TE8) protein. (A) Human IL-10 PDB 1ILK. (B) Homology model of TeHV-3 IL-10 constructed from 1ILK with suboptimal residues shown. (C) Ramachandran plot of homology model used to identify suboptimal residues. (D) Human IL-10 PDB 1ILK (blue) superposed with the homology model of TeHV-3 IL-10 (red).

At necropsy, few macroscopic lesions were observed. A single tortoise (4295/41 D) exhibited diphtheritic membranes on the buccal and esophageal mucosa. No other lesions were observed in the infected tortoises or the mock-infected tortoise. Histopathological analysis of various organs (lung, spleen, brain, kidney, and liver) of the infected tortoises revealed mild lesions consistent with previous reports (8, 13) (Fig. 9). These lesions consisted of interstitial heterophilic infiltration of the lung, congestion and heterophilic infiltration of the red pulp of the spleen, and glial cell infiltration of the brain (telencephalon). The kidney and the liver did not show significant histopathological modifications. Eosinophilic intranuclear inclusion bodies were not detected in any of the samples examined.

TABLE 3 Best achievable root mean square deviation (RMSD) values for main chain superposition between homology models and the TeHV-3 TE7 (SEMA) and TE8 (v-IL-10) proteins

TeHV-3 protein	PDB accession no.	Model	RMSD (Å)
TE7	1Q47	Mouse SEMA-3A	1.218
	1OLZ	Human SEMA-4D	1.257
	3NVQ	Human SEMA-7A	1.358
	3OKW	Mouse SEMA-6A	1.496
	3NVX	Vaccinia virus SEMA	2.655
	3AL9	Mouse plexin A2	3.911
	3NVN	Ectromelia virus SEMA	4.347
TE8	1ILK	Human IL-10	1.040
	1LQS	HHV-5 IL-10	1.300

To investigate tissue tropism, viral load was estimated in various organs (lung, spleen, brain, kidney, and liver) using a qPCR method developed for the purpose. Consistent with earlier reports (8, 13, 16), the virus was detected in all organs tested (Fig. 10A). However, in all three infected tortoises, the highest viral load was detected in the brain ($P < 0.05$), slightly greater than the viral load in the spleen. The other organs revealed lower and comparable viral loads ($P < 0.01$). At 3-week intervals, peripheral mononuclear blood cells were collected during the entire course of the experiment and subjected to qPCR analyses. TeHV-3 was not detected in any of the samples (data not shown).

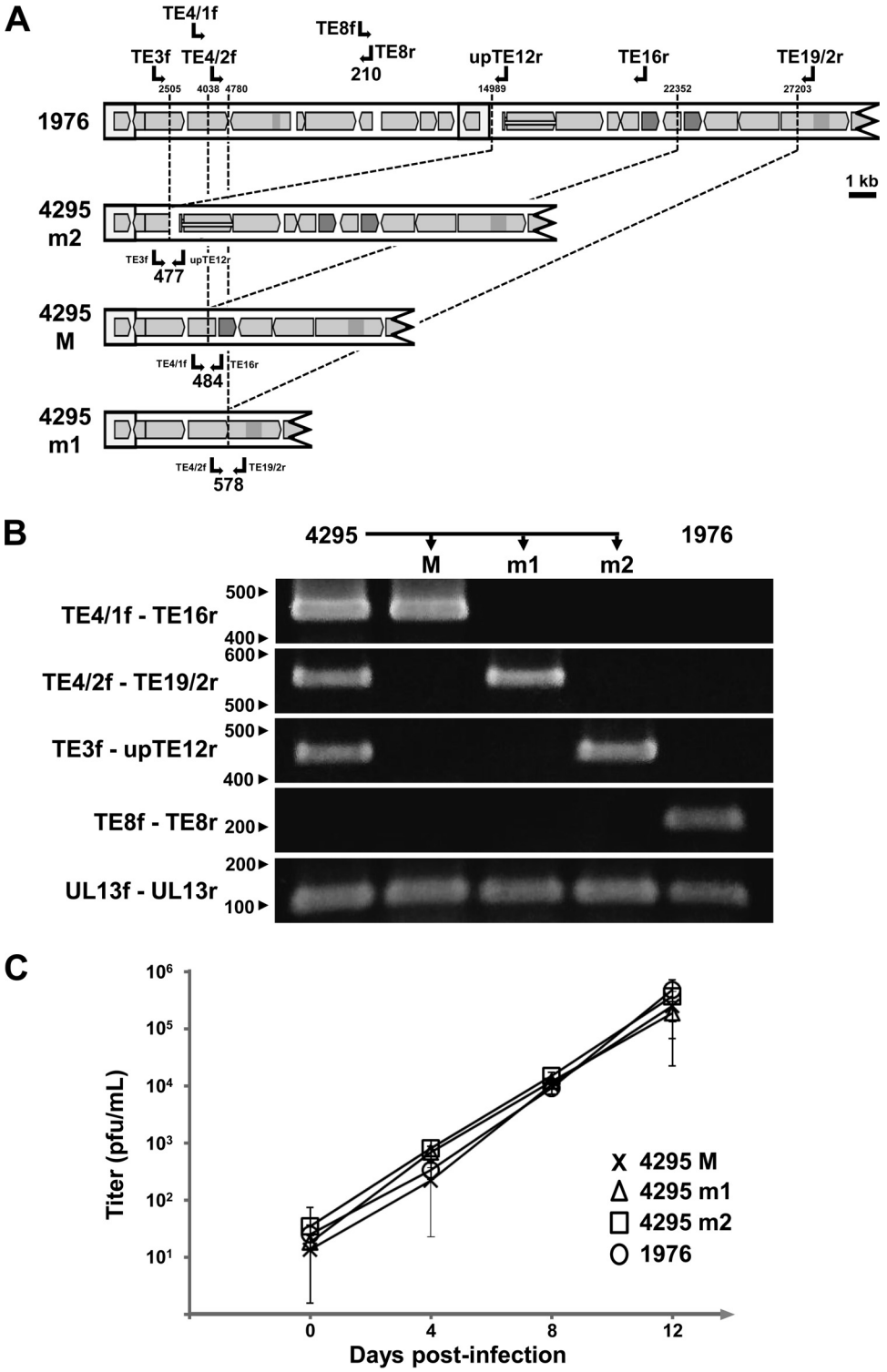

FIG 7 Effects of the deletions in the m1, m2, and M forms of strain 4295 on viral growth *in vitro*. (A) Schematic representation of the regions in the strain 1976 genome corresponding to the deletions in the three forms of strain 4295, with the coordinates of the deletions indicated above the strain 1976 genome. Arrows represent primers (Table 1) designed for PCR amplification of the regions containing the deletions. Amplicon sizes are indicated in bold below the genomes of the strain 4295 subclones. PCR amplification performed with the UL13f and UL13r primers led to a product of 106 bp. (B) Characterization by PCR of the strain 4295 subclones representing the three genome forms. Strain 4295 prior to subcloning (containing the three forms) and strain 1976 were used as controls. The positions of markers (in base pairs) are marked by arrowheads. (C) Multistep growth curves of the strain 4295 subclones and strain 1976.

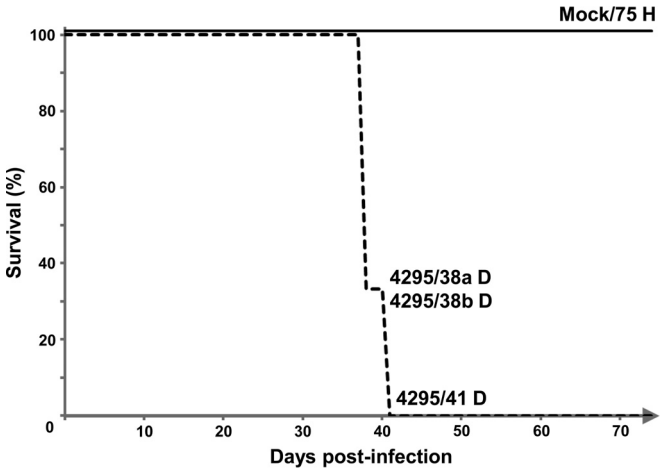

FIG 8 Pathogenesis of strain 4295. On day 0, tortoises ($n = 4$) aged 5.4 years (mean weight ± SD, 219.5 g ± 53.1 g) were infected ($n = 3$) with strain 4295 (consisting of a mixture of three deletion mutants) or mock-infected ($n = 1$). Percentage survival is expressed according to the days postinfection. The tortoises were named according to the following scheme: the inoculation performed (strain 4295 or mock)/the time postinfection at which death occurred (when more than one tortoise died on the same day, they were further ranked by the addition of a lowercase letter) and an uppercase letter to describe the clinical state before euthanasia (D, diseased; H, healthy).

Finally, we used the PCR assays described above to determine whether the viral load detected in the various organs in Fig. 10A represented all three forms of 4295 (Fig. 10B). The m1 form was not detected in any sample. The m2 form was detected with the highest frequency and was identified in all organs shown to be positive for TeHV-3 by PCR of a gene present in all forms (UL13). Like the m2 form, the M form was detected in the brain of all infected subjects. However, its presence in the other organs was reduced compared to the m2 form. The M form was detected in only a fraction of the lungs and the spleens positive for the m2 form, and with the exception of one sample, it was not detected in the kidney and liver samples that were positive for the m2 form.

As the PCR data presented above were derived from tortoises that had been coinfected with a mixture of the three forms present in strain 4295, they should be interpreted cautiously in terms of the fitness of individual forms. Indeed, even if it might be viewed as unlikely, *in vivo* interactions between the forms, such as helper effects resulting from superinfection of the same host cells or secretion of soluble factors, cannot be excluded. Consequently, the conclusions described below will need to be confirmed by experimental infection of animals by the individual genotypes. The ability of the m2 form to spread throughout the body suggests that the region of the TeHV-3 genome encompassing TE3 to TE11 genes (corresponding to the deletion in this form) is not essential for virulence and neuroinvasiveness. Similarly, the ability of the M form to invade the brain suggests that the region carrying TE4 to TE16 genes is not essential for neuroinvasiveness. However, compared to the m2 form, the restricted tropism of the M form for the other organs suggests that one or more genes in the region containing TE12 to TE16 genes may contribute to viral spread *in vivo*.

Setting aside any possible interactions that may have occurred between the three forms *in vivo*, the results of the analyses presented in Fig. 10B suggest that the region encompassing TE4 to TE11 (the deletion common to all three forms), as well as not being required for viral growth *in vitro*, is not essential for virulence of TeHV-3 *in vivo*. Importantly, the inability of the m1 form to spread *in vivo* (despite coinfection with the two other genotypes) demonstrates that the region encompassing UL55B to TE19 genes contains one or more genes that are essential (alone or in combination) for viral spread *in vivo*. As the m1 form was shown to grow *in vitro* as efficiently as the two other forms (m2 and M) and strain 1976, the results of the present study encourage the

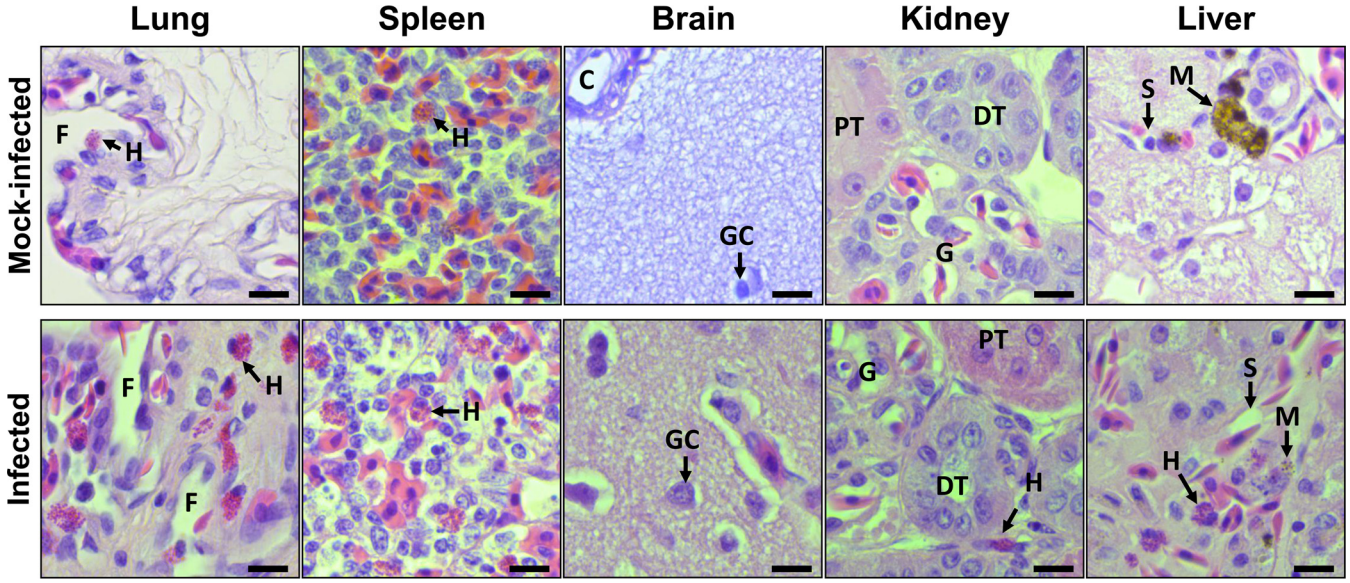

FIG 9 Histopathological characterization of the lesions induced by strain 4295 (consisting of a mixture of three deletion mutants). The indicated organs were collected from all tortoises at the end of the experiment described in the legend to Fig. 8 and were processed for histological examination. The images were collected from tortoises Mock/75 H and 4295/41 D, the latter having been selected as representative of the infected group. H, heterophil; F, faveolae; GC, glial cell; C, capillary; PT, proximal tubule; DT, distal tubule; G, glomerulus; S, sinusoid; and M, melanomacrophage. Bars, 20 μm.

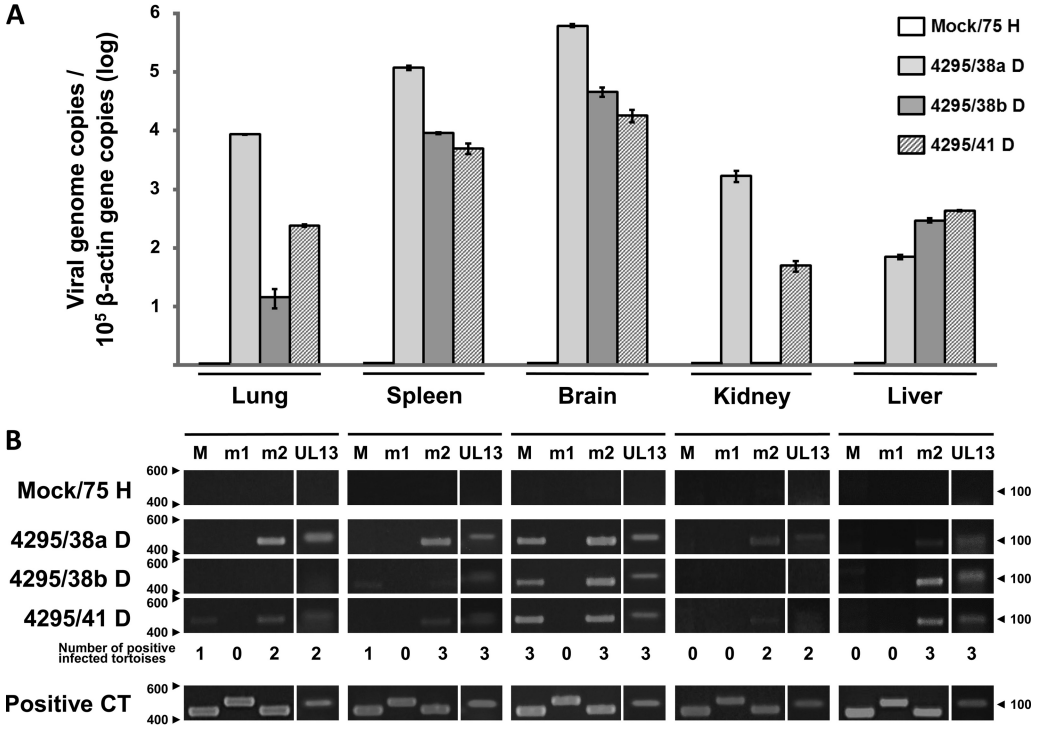

FIG 10 Tissue tropism of strain 4295. DNA was extracted from the indicated organs of all tortoises at the end of the experiment described in the legend of Fig. 8. (A) Analysis of viral gene copy number by qPCR. Individual values represent the means from triplicate measurements ± SD. (B) Analysis of the presence of the three forms present in strain 4295, using primers specific for each form (m1, m2, and M) and all three forms (UL13). See Fig. 7A and B and Table 1 for details of the primers. Strain 4295 (containing the three forms) was used as positive (CT) control.

testing of the m1 subclone of strain 4295 as an attenuated vaccine candidate.

While this study was under review, a version of the strain 1976 genome sequence became available (59). In comparison with our sequence, this contains a 9,521-nt deletion (nt 5469 to 14989) starting in TE5 and ending near the left end of U_L and thus lacks seven genes, including TE7 and TE8. The similarity in location between this deletion and those in strain 4295 indicates that the TeHV-3 genome has a propensity for losing information in this region. Based on the pathogenic properties of strain 4295, we predict that the deletion in strain 1976 will not affect viral spread *in vivo*. However, the sequence also has frameshifts in TE19 that, if not due to error, may affect virulence, since this gene is absent from the m1 form of strain 4295.

Our study reports the first complete genome sequences of TeHV-3 strains 1976 and 4295, the latter comprising a mixture of three deletion mutants that were sequenced as the mixture and as subclones. The sequence of strain 1976 revealed a novel genome structure in the family *Herpesviridae*. Genetic analysis highlighted the presence of genes related to cellular genes (SEMA and IL-10) that have not been reported previously in alphaherpesviruses. Phylogenetic analysis showed that TeHV-3 is most closely related to turtle herpesviruses and suggested the classification of this virus in the genus *Scutavirus*. Importantly, *in vitro* and *in vivo* analyses demonstrated that the TeHV-3 genome contains large regions that are essential neither for viral replication *in vitro* nor for virulence *in vivo*. They also indicated that spread of the m1 form of strain 4295 is attenuated *in vivo*, thus indicating that this form would be a good starting point for the development of vaccine candidates. In conclusion, the present study represents a major step toward the characterization of an important viral pathogen of tortoises and the development of effective prophylactic measures against TeHV-3 disease.

ACKNOWLEDGMENTS

This work was supported by the University of Liège, the Belgian Science Policy (Belspo) (BELVIR IAP7/45), the Fonds National Belge de la Recherche Scientifique (FNRS), and the United Kingdom Medical Research Council (grant number MC_UU_12014/3). B.G.D. is a research fellow of the FNRS.

We thank Wai Kwong Lee and Andrew Carswell (BHF Glasgow Cardiovascular Research Centre, University of Glasgow) for providing Sanger DNA sequencing services.

REFERENCES

1. **Pellett PE, Davison AJ, Eberle R, Ehlers B, Hayward GS, Lacoste V, Minson AC, Nicholas J, Roizman B, Studdert MJ, Wang F.** 2012. Order—Herpesvirales, p 99–107. *In* King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy. Elsevier, San Diego, CA. http://dx.doi.org/10.1016/B978-0-12-384684-6.00005-7.
2. **Pellett PE, Roizman B.** 2013. Herpesviridae, p 1802–1822. *In* Knipe DM, Howley PM (ed), Fields virology, 6th ed, vol 2. Lippincott Williams & Wilkins, Philadelphia, PA.
3. **Bicknese EJ, Childress AL, Wellehan JF, Jr.** 2010. A novel herpesvirus of the proposed genus Chelonivirus from an asymptomatic bowsprit tortoise (*Chersina angulata*). J Zoo Wildl Med **41**:353–358. http://dx.doi.org/10.1638/2009-0214R.1.
4. **Ackermann M, Koriabine M, Hartmann-Fritsch F, de Jong PJ, Lewis TD, Schetle N, Work TM, Dagenais J, Balazs GH, Leong JA.** 2012. The genome of Chelonid herpesvirus 5 harbors atypical genes. PLoS One **7**:e46623. http://dx.doi.org/10.1371/journal.pone.0046623.

5. **Origgi FC.** 2006. Herpesvirus in tortoises, p 814–821. *In* Mader DR (ed), Reptile medicine and surgery, 2nd ed. Saunders, St. Louis, MO.
6. **Marschang RE, Gleiser CB, Papp T, Pfitzner AJ, Bohm R, Roth BN.** 2006. Comparison of 11 herpesvirus isolates from tortoises using partial sequences from three conserved genes. Vet Microbiol **117:**258–266. http://dx.doi.org/10.1016/j.vetmic.2006.06.009.
7. **Origgi FC, Klein PA, Mathes K, Blahak S, Marschang RE, Tucker SJ, Jacobson ER.** 2001. Enzyme-linked immunosorbent assay for detecting herpesvirus exposure in Mediterranean tortoises (spur-thighed tortoise [Testudo graeca] and Hermann's tortoise [Testudo hermanni]). J Clin Microbiol **39:**3156–3163. http://dx.doi.org/10.1128/JCM.39.9.3156-3163.2001.
8. **Origgi FC, Romero CH, Bloom DC, Klein PA, Gaskin JM, Tucker SJ, Jacobson ER.** 2004. Experimental transmission of a herpesvirus in Greek tortoises (Testudo graeca). Vet Pathol **41:**50–61. http://dx.doi.org/10.1354/vp.41-1-50.
9. **Marschang RE, Gravendyck M, Kaleta EF.** 1997. Herpesviruses in tortoises: investigations into virus isolation and the treatment of viral stomatitis in Testudo hermanni and T. graeca. Zentralbl Veterinarmed B **44:**385–394. http://dx.doi.org/10.1111/j.1439-0442.1997.tb01123.x.
10. **Muro J, Ramis A, Pastor J, Velarde R, Tarres J, Lavin S.** 1998. Chronic rhinitis associated with herpesviral infection in captive spur-thighed tortoises from Spain. J Wildl Dis **34:**487–495. http://dx.doi.org/10.7589/0090-3558-34.3.487.
11. **Martel A, Blahak S, Vissenaekens H, Pasmans F.** 2009. Reintroduction of clinically healthy tortoises: the herpesvirus Trojan horse. J Wildl Dis **45:**218–220. http://dx.doi.org/10.7589/0090-3558-45.1.218.
12. **Zapata AG, Varas A, Torroba M.** 1992. Seasonal variations in the immune system of lower vertebrates. Immunol Today **13:**142–147. http://dx.doi.org/10.1016/0167-5699(92)90112-K.
13. **Teifke JP, Lohr CV, Marschang RE, Osterrieder N, Posthaus H.** 2000. Detection of chelonid herpesvirus DNA by nonradioactive in situ hybridization in tissues from tortoises suffering from stomatitis-rhinitis complex in Europe and North America. Vet Pathol **37:**377–385. http://dx.doi.org/10.1354/vp.37-5-377.
14. **Une Y, Uemura K, Nakano Y, Kamiie J, Ishibashi T, Nomura Y.** 1999. Herpesvirus infection in tortoises (Malacochersus tornieri and Testudo horsfieldii). Vet Pathol **36:**624–627. http://dx.doi.org/10.1354/vp.36-6-624.
15. **Hervas J, Sanchez-Cordon PJ, de Chacon Lara F, Carrasco L, Gomez-Villamandos JC.** 2002. Hepatitis associated with herpes viral infection in the tortoise (Testudo horsfieldii). J Vet Med B Infect Dis Vet Public Health **49:**111–114. http://dx.doi.org/10.1046/j.1439-0450.2002.00522.x.
16. **Marschang RE, Milde K, Bellavista M.** 2001. Virus isolation and vaccination of Mediterranean tortoises against a chelonid herpesvirus in a chronically infected population in Italy. Dtsch Tierarztl Wochenschr **108:**376–379.
17. **Blahak S, Steiner M.** 2010. Epidemiological investigations into the persistence of antibodies against tortoise herpesvirus strains in the species *T. graeca*, *T. hermanni* and *T. horsfieldii*, abstr 1st Int Conf Reptile Amphibian Med, Munich, 4 March 2010. Verlag, Munich, Germany.
18. **Marschang RE, Frost JW, Gravendyck M, Kaleta EF.** 2001. Comparison of 16 chelonid herpesviruses by virus neutralization tests and restriction endonuclease digestion of viral DNA. J Vet Med B Infect Dis Vet Public Health **48:**393–399. http://dx.doi.org/10.1046/j.1439-0450.2001.00450.x.
19. **Clark HF, Karzon DT.** 1967. Terrapene heart (TH-1), a continuous cell line from the heart of the box turtle Terrapene carolina. Exp Cell Res **48:**263–268. http://dx.doi.org/10.1016/0014-4827(67)90351-5.
20. **Lete C, Palmeira L, Leroy B, Mast J, Machiels B, Wattiez R, Vanderplasschen A, Gillet L.** 2012. Proteomic characterization of bovine herpesvirus 4 extracellular virions. J Virol **86:**11567–11580. http://dx.doi.org/10.1128/JVI.00456-12.
21. **Markine-Goriaynoff N, Georgin JP, Goltz M, Zimmermann W, Broll H, Wamwayi HM, Pastoret PP, Sharp PM, Vanderplasschen A.** 2003. The core 2 beta-1,6-N-acetylglucosaminyltransferase-mucin encoded by bovine herpesvirus 4 was acquired from an ancestor of the African buffalo. J Virol **77:**1784–1792. http://dx.doi.org/10.1128/JVI.77.3.1784-1792.2003.
22. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.
23. **Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I.** 2009. ABySS: a parallel assembler for short read sequence data. Genome Res **19:**1117–1123. http://dx.doi.org/10.1101/gr.089532.108.
24. **Ewing B, Hillier L, Wendl MC, Green P.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res **8:**175–185.
25. **Ewing B, Green P.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res **8:**186–194.
26. **Tsai IJ, Otto TD, Berriman M.** 2010. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. Genome Biol **11:**R41. http://dx.doi.org/10.1186/gb-2010-11-4-r41.
27. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics **26:**589–595. http://dx.doi.org/10.1093/bioinformatics/btp698.
28. **Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D.** 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform **14:**193–202. http://dx.doi.org/10.1093/bib/bbs012.
29. **Gatherer D, Seirafian S, Cunningham C, Holton M, Dargan DJ, Baluchova K, Hector RD, Galbraith J, Herzyk P, Wilkinson GW, Davison AJ.** 2011. High-resolution human cytomegalovirus transcriptome. Proc Natl Acad Sci U S A **108:**19755–19760. http://dx.doi.org/10.1073/pnas.1115861108.
30. **Wilkie GS, Davison AJ, Watson M, Kerr K, Sanderson S, Bouts T, Steinbach F, Dastjerdi A.** 2013. Complete genome sequences of elephant endotheliotropic herpesviruses 1A and 1B determined directly from fatal cases. J Virol **87:**6700–6712. http://dx.doi.org/10.1128/JVI.00655-13.
31. **Costes B, Fournier G, Michel B, Delforge C, Raj VS, Dewals B, Gillet L, Drion P, Body A, Schynts F, Lieffrig F, Vanderplasschen A.** 2008. Cloning of the koi herpesvirus genome as an infectious bacterial artificial chromosome demonstrates that disruption of the thymidine kinase locus induces partial attenuation in Cyprinus carpio koi. J Virol **82:**4955–4964. http://dx.doi.org/10.1128/JVI.00211-08.
32. **Costes B, Raj VS, Michel B, Fournier G, Thirion M, Gillet L, Mast J, Lieffrig F, Bremont M, Vanderplasschen A.** 2009. The major portal of entry of koi herpesvirus in Cyprinus carpio is the skin. J Virol **83:**2819–2830. http://dx.doi.org/10.1128/JVI.02305-08.
33. **Vanderplasschen A, Bublot M, Dubuisson J, Pastoret P-P, Thiry E.** 1993. Attachment of the gammaherpesvirus bovine herpesvirus 4 is mediated by the interaction of gp8 glycoprotein with heparinlike moieties on the cell surface. Virology **196:**232–240. http://dx.doi.org/10.1006/viro.1993.1471.
34. **Dewals B, Boudry C, Gillet L, Markine-Goriaynoff N, de Leval L, Haig DM, Vanderplasschen A.** 2006. Cloning of the genome of Alcelaphine herpesvirus 1 as an infectious and pathogenic bacterial artificial chromosome. J Gen Virol **87:**509–517. http://dx.doi.org/10.1099/vir.0.81465-0.
35. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797. http://dx.doi.org/10.1093/nar/gkh340.
36. **Kumar S, Nei M, Dudley J, Tamura K.** 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform **9:**299–306. http://dx.doi.org/10.1093/bib/bbn017.
37. **Drummond AJ, Suchard MA, Xie D, Rambaut A.** 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol **29:**1969–1973. http://dx.doi.org/10.1093/molbev/mss075.
38. **Le SQ, Gascuel O.** 2008. An improved general amino acid replacement matrix. Mol Biol Evol **25:**1307–1320. http://dx.doi.org/10.1093/molbev/msn067.
39. **Drummond AJ, Ho SY, Phillips MJ, Rambaut A.** 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol **4:**e88. http://dx.doi.org/10.1371/journal.pbio.0040088.
40. **Gernhard T.** 2008. The conditioned reconstructed process. J Theor Biol **253:**769–778. http://dx.doi.org/10.1016/j.jtbi.2008.04.005.
41. **Gatherer D.** 2007. TreeAdder: a tool to assist the optimal positioning of a new leaf into an existing phylogenetic tree. Open Bioinfo J **1:**1–2. (Letter.) http://dx.doi.org/10.2174/1875036200701010001.
42. **Ouyang P, Rakus K, van Beurden SJ, Westphal AH, Davison AJ, Gatherer D, Vanderplasschen AF.** 2014. IL-10 encoded by viruses: a remarkable example of independent acquisition of a cellular gene by viruses and its subsequent evolution in the viral genome. J Gen Virol **95:**245–262. http://dx.doi.org/10.1099/vir.0.058966-0.
43. **Schmidt HA, Strimmer K, Vingron M, von Haeseler A.** 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18:**502–504. http://dx.doi.org/10.1093/bioinformatics/18.3.502.
44. **Shimodaira H, Hasegawa M.** 2001. CONSEL: for assessing the confidence

of phylogenetic tree selection. Bioinformatics **17:**1246–1247. http://dx.doi.org/10.1093/bioinformatics/17.12.1246.
45. **Levitt M.** 1992. Accurate modeling of protein conformation by automatic segment matching. J Mol Biol **226:**507–533. http://dx.doi.org/10.1016/0022-2836(92)90964-L.
46. **Fechteler T, Dengler U, Schomburg D.** 1995. Prediction of protein three-dimensional structures in insertion and deletion regions: a procedure for searching data bases of representative protein fragments using geometric scoring criteria. J Mol Biol **253:**114–131. http://dx.doi.org/10.1006/jmbi.1995.0540.
47. **Wang J, Cieplak P, Kollman PA.** 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem **21:**1049–1074.
48. **Labute P.** 2008. The generalized Born/volume integral implicit solvent model: estimation of the free energy of hydration using London dispersion instead of atomic surface area. J Comput Chem **29:**1693–1698. http://dx.doi.org/10.1002/jcc.20933.
49. **Ramachandran GN, Ramakrishnan C, Sasisekharan V.** 1963. Stereochemistry of polypeptide chain configurations. J Mol Biol **7:**95–99. http://dx.doi.org/10.1016/S0022-2836(63)80023-6.
50. **Telford EA, Watson MS, McBride K, Davison AJ.** 1992. The DNA sequence of equine herpesvirus-1. Virology **189:**304–316. http://dx.doi.org/10.1016/0042-6822(92)90706-U.
51. **Telford EA, Watson MS, Perry J, Cullinane AA, Davison AJ.** 1998. The DNA sequence of equine herpesvirus-4. J Gen Virol **79**(Part 5)**:**1197–1203.
52. **Stow ND, Davison AJ.** 1986. Identification of a varicella-zoster virus origin of DNA replication and its activation by herpes simplex virus type 1 gene products. J Gen Virol **67**(Part 8)**:**1613–1623.
53. **Koff A, Tegtmeyer P.** 1988. Characterization of major recognition sequences for a herpes simplex virus type 1 origin-binding protein. J Virol **62:**4096–4103.
54. **Weir HM, Stow ND.** 1990. Two binding sites for the herpes simplex virus type 1 UL9 protein are required for efficient activity of the oriS replication origin. J Gen Virol **71**(Part 6)**:**1379–1385.
55. **Davison AJ.** 2010. Herpesvirus systematics. Vet Microbiol **143:**52–69. http://dx.doi.org/10.1016/j.vetmic.2010.02.014.
56. **Myster F, Palmeira L, Sorel O, Bouillenne F, DePauw E, Schwartz-Cornil I, Vanderplasschen A, Dewals BG.** 2015. Viral semaphorin inhibits dendritic cell phagocytosis and migration but is not essential for gammaherpesvirus-induced lymphoproliferation in malignant catarrhal fever. J Virol **89:**3630–3647. http://dx.doi.org/10.1128/JVI.03634-14.
57. **Liu H, Juo ZS, Shim AH, Focia PJ, Chen X, Garcia KC, He X.** 2010. Structural basis of semaphorin-plexin recognition and viral mimicry from Sema7A and A39R complexes with PlexinC1. Cell **142:**749–761. http://dx.doi.org/10.1016/j.cell.2010.07.040.
58. **Pasmans F, Blahak S, Martel A, Pantchev N.** 2008. Introducing reptiles into a captive collection: the role of the veterinarian. Vet J **175:**53–68. http://dx.doi.org/10.1016/j.tvjl.2006.12.009.
59. **Origgi FC, Tecilla M, Pilo P, Aloisio F, Otten P, Aguilar-Bultet L, Sattler U, Roccabianca P, Romero CH, Bloom DC, Jacobson ER.** 2015. A genomic approach to unravel host-pathogen interaction in chelonians: the example of testudinid herpesvirus 3. PLoS One **10:**e0134897. http://dx.doi.org/10.1371/journal.pone.0134897.