ORIGINAL ARTICLE

Comparative sequence and structural analysis of the ORF095 gene, a vaccinia virus A4L homolog of capripoxvirus in sheep and goats

Aparna Madhavan¹ · Gnanavel Venkatesan1 · Amit Kumar1 · Sargam Arya1 · A. B. Pandey1

Received: 3 November 2019 / Accepted: 23 March 2020 / Published online: 19 April 2020 © Springer-Verlag GmbH Austria, part of Springer Nature 2020

Abstract

Sheeppox and goatpox are important transboundary animal viral diseases of sheep and goats caused by sheeppox virus (SPPV) and goatpox virus (GTPV), respectively, of the genus *Capripoxvirus*, family *Poxviridae*. Among the proteins encoded by the capripoxvirus (CaPV) genome, ORF095 (vaccinia virus A4L homolog) is an immunodominant virion core protein that plays a pivotal role in virus assembly and morphogenesis. In the present study, sequence analysis of the ORF095 genes of 27 SPPV and GTPV isolates or feld samples from diferent geographical regions of India was performed, and structure was prediction was done by homology modeling. A multiple sequence alignment of diferent CaPV isolates revealed that CaPV-A4L is highly conserved, with several species-specifc signature residues, namely A93, A216, A315, G136 and G146 in GTPV, G47, A63, A168 and A276 in SPPV, and G48 and C98 in lumpy skin disease virus (LSDV). Phylogenetically, the CaPV isolates were separated into three major clusters, GTPV, SPPV and LSDV, based on the complete coding sequence of the CaPV-A4L gene. Genus-specifc clustering of poxviruses was observed in phylogenetic analysis based on A4L protein homologs of chordopoxviruses. A secondary structure prediction showed the presence of six α-helices and one β-sheet as well as some coils. The signature residues identifed here are potentially useful for genotyping, and the predicted characteristics of the CaPV-A4L protein make it an ideal candidate for use as an immunogenic or diagnostic antigen for the development of immunoassays in the sero-evaluation of CaPV in target hosts.

Introduction

Capripox infection is a transboundary and World Health Organization for Animal Health (OIE- notifable disease of sheep, goats and cattle caused by sheeppox virus (SPPV), goatpox virus (GTPV) and lumpy skin disease virus (LSDV) virus, respectively, of the genus *Capripox virus*, family *Poxviridae* [\[1](#page-9-0)]. The geographical range of sheeppox and goatpox includes Africa (north of the equator), the Middle East, Iran, Iraq, Bangladesh, Afghanistan, Pakistan, China, India and European countries adjoining Middle-East regions, namely Turkey, Greece and Bulgaria [[2\]](#page-9-1). Recent outbreaks in Vietnam, Mongolia, Morocco and Azerbaijan indicate the emergence of capripoxviruses (CaPVs) in previously unafected regions [[3](#page-11-0)]. These outbreaks pose a signifcant economic

Handling Editor: William G Dundon.

 \boxtimes Gnanavel Venkatesan gnanamvirol@gmail.com threat due to productivity losses, high mortality, hide damage, and restrictions on world trade [[4,](#page-11-1) [5](#page-11-2)]. CAPVs cause a highly contagious disease characterized by fever, oculonasal discharge, and pock lesions on the skin and mucosae of the respiratory and gastrointestinal tract [[6\]](#page-11-3). SPPV and GTPV have host preference for sheep and goats, respectively, and are considered diferent entities that provide only partial cross-protection [[7](#page-11-4)] although they have a very close antigenic relationship [\[8](#page-11-5)]. Apart from conventional diagnostic approaches, PCR-based diagnostic techniques and genome sequencing have provided sensitive and powerful techniques for the identifcation and diferentiation of SPPV and GTPV [\[3](#page-11-0), [8](#page-11-5)]. It is possible to distinguish CaPVs at the genetic level by sequence comparison and phylogenetic analysis of individual genes or the whole genome sequence and also by polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP) analysis [\[9](#page-11-6), [10](#page-11-7)]. Multigene analysis is more reliable than analysis of single genes for determining genetic relatedness of because of their large genome size. SPPV and GTPV can cause cross-infection, and virus identifcation based on the host animal species from which the virus was obtained is not valid. Recent developments in molecular biology have provided deep insights into the

Division of Virology, ICAR-Indian Veterinary Research Institute, Mukteswar 263 138, Nainital District, Uttarakhand, India

cross-species transmission of SPPV and GTPV that makes it necessary to characterize the CaPV isolates from sheep and goats circulating at the feld level [[7\]](#page-11-4).

The present study focused on genetic and structural analysis of the immunodominant virion core protein gene ORF095, a vaccinia virus A4L homolog [[11](#page-11-8), [12\]](#page-11-9), using SPPV and GTPV isolates and clinical samples $(n = 27)$. The vaccinia virus A4L protein encodes a highly conserved 39-kDa immunodominant virion core protein that plays a role in assembly and disassembly of the virion [\[13\]](#page-11-10). To date, genetic characterization of CaPV-A4L has been limited to some reports on expression in a prokaryotic host [\[14](#page-11-11)]. To our knowledge, this is the frst study to analyze the full-length A4L gene of CaPVs at the sequence and structural levels using SPPV and GTPV isolates from India and elsewhere. A comparative analysis of A4L homologs of poxviruses was done to assess their evolutionary relationship.

Materials and methods

Virus isolates, clinical samples, and primers

A total of twenty cell culture virus isolates and seven feld clinical samples of GTPV and SPPV maintained in the virus repository of the poxvirus laboratory of the Division of Virology, Indian Veterinary Research Institute, Mukteswar, were included in this study. All of the virus isolates and samples used in this study were originally collected from animals with natural sheep pox or goat pox infections at various geographical locations in India and at different times. Details of the virus isolates and samples used in this study and other isolates of CaPV, including LSDV isolates, used for sequence analysis are listed in Table [1.](#page-2-0) All SPPV and GTPV isolates were propagated in Vero cells obtained from ATCC (CCL81) using Eagle's minimum essential medium (EMEM) along with 2% bovine calf serum (BCS) as maintenance serum. Infected cells were harvested when 80% of the cells showed a cytopathic efect (CPE) and used for extraction of total viral genomic DNA (gDNA) using a QIAamp DNA Mini Kit (QIAGEN, USA) as per the manufacturer's protocol. The clinical samples were processed as 10% homogenates in PBS and then used for gDNA extraction. Primers for amplifcation of the full-length A4L genes of CaPVs were designed based on the complete genome sequences of GTPV, SPPV and LSDV strains available in the GenBank database and synthesized commercially.

PCR amplifcation, cloning and sequencing of the CaPV‑A4L gene

The extracted CaPV gDNA was used for PCR amplifcation of the A4L gene, using standard cycling conditions and an optimized annealing temperature using in-house-designed primers (ORF095 fwd, 5'-GGCCATGGCGATGGACTT CATGAAAAAATATAC-3'; ORF095 rev, 5'-GGAAGC TTTTTGCTGTTATTATCATCTAG-3') covering the full length of the gene. In brief, PCR amplifcation was carried out using 1X GoTaq High Fidelity PCR Master Mix (Promega, USA), 10 pmol of each primer and 1 µL of template DNA (\sim 50-100 ng) under optimized cycling parameters: 95 °C for 5 min, 35 cycles of denaturation (95 °C for 45 s), annealing (53 °C for 45 sec) and extension (72 °C for 1 min), followed by a final extension step at $72 \degree C$ for 10 min. The PCR product was then purifed using a gel purifcation kit (QIAGEN, MD, USA) using the manufacturer's standard protocol and ligated into pGEM®-T Easy Vector (Promega, Madison, WI, USA) using a T/A cloning strategy (data not shown). These ligated products then were then used to transform *E. coli* Top 10F' cells by the heat shock method. Positive recombinant clones were selected using antibiotics and a blue-white colony system and sent for commercial sequencing (Delhi University South Campus, New Delhi, India). Clones were selected and sequenced twice before genetic analysis. The complete nucleotide sequences of the A4L genes of diferent SPPV and GTPV isolates used in this study were submitted at GenBank, and their accession numbers are listed in Table [1.](#page-2-0)

Multiple sequence alignment and phylogenetic analysis

A multiple sequence alignment of CaPV-A4L gene sequences of Indian CaPV isolates from this study and those of other CaPVs retrieved from the GenBank database was generated using the Clustal W program [[15\]](#page-11-12). Percent identity values were determined using the MegAlign program of the DNASTAR package (Lasergene 6.0, DNASTAR Inc., Madison, USA). Phylogenetic trees based on the nt and the deduced aa sequences of the A4L genes of CaPVs (Table [1\)](#page-2-0) and other animal poxviruses (Table [2\)](#page-3-0) were constructed by the neighbor-joining method in MEGA 7.0 software, using 1,000 bootstrap replicates [[16\]](#page-11-13).

Protein modeling and structural features of CaPV‑A4L

The consensus amino acid sequence of the CaPV-A4L protein was analyzed using various protein structure prediction programs, including Protean (DNASTAR), PSIPRED [\[17](#page-11-14)], and coils2 [[18\]](#page-11-15) for secondary structure prediction and identifcation of other protein characteristics. A structural model of CaPV-A4L was made using the I-TASSER online server, and the accuracy and quality of the predicted structure were estimated based on C-score, TM-score, and RMSD values [[19,](#page-11-16) [20\]](#page-11-17).

Table 2 List of representative poxviruses with A4L gene sequences used in comparative phylogenetic analysis

Virus	ORF (aa	A4L gene/locus designation	Position	Accession no.			
	length)			Nucleotide	Protein		
Genus: Orthopoxvirus							
Cowpox virus (CPXV-BR)	295	CPXV-136	130026-130913	AF482758	AAM13577		
Racoonpox virus (RCNV-Herman)	289	ACG19_gp119	119616-120485	NC_027213	YP_009143431		
Rabbitpox virus (RPXV-Utr)	283	RPXV-Utr-112	118062-118913	AY484669	AAS49825		
Taterapox virus (TATV-DAH68	282	TATV-DAH68-125	113140-113988	DQ437594	ABD97691		
Vaccinia virus (VACV-WR)	281	VACV-WR-123	113281-114126	AY243312	AAO89402		
Monkeypox virus (MPXV-COG03)	281	MPXV_RCG2003_358_120	113474-114319	DQ011154	AAY97114		
Horsepox virus (HSPV-MNR76)	281	HSPV124	124632-125477	DQ792504	ABH08230		
Ectromelia virus (ECTV-Mos)	281	EVM107	122096-122941	AF012825	AAM92411		
Camelpox virus (CMLV-CMS)	275	CMP-120L	114356-115183	AY009089	AAG37604		
Variola virus (VARV-IND64)	271	VARV-IND64_vel4-112	104738-105553	DQ437585	ABG44290		
Genus: Centapoxvirus							
Yokavirus (YKV-DakArB)	238	YKV105c	98017-98733	HQ849551	AEN03694		
Genus: Capripoxvirus							
Sheeppox virus (SPPV-TU)	161	SPPV_91	88850-89335	NC004002	NP_659667		
Goatppox virus (GTPV-Pellor)	161	GTPV_gp091	88714-89199	NC004003	YP_001293286		
Lumpy skin disease virus (LSDV-LW1959)	161	LW095	89162-89647	AF409138	AAN02820.1		
Genus: Suipoxvirus							
Swinepox virus (SWPV-Neb)	154	SPV092	87766-88230	AF410153	AAL69831		
Genus: Leporipoxvirus							
Myxoma virus (MYXV-Lau)	159	MYXV_gp097	91959-92438	NC_001132	NP_051807		
Rabbit fibroma virus (RFV-Kas)	153	SFV_s093L	91166-91627	NC001266	NP_051982		
Genus: Parapoxvirus							
Orf virus (ORFV-SA00)	328	ORFVgORF080	83600-84586	NC_005336	NP_957857		
Parapoxvirus of red deer (PRDV-HL)	267	SB87_gp078	86903-87706	NC_025963	YP_009112819		
Pseudocowpox virus (PPXV-VR634)	262	PCPV_gp083	91655-92443	NC_013804	YP_003457386		
Bovine popular stomatitis virus (BPSV-AR02)	244	BPSVgORF080	83680-84414	NC_005337	NP_957989		
Genus: Avipoxvirus							
Fowlpox virus (FWPV)	243	FPV168	201565-202431	NC_002188	NP_039131		
Genus: <i>Molluscipoxvirus</i>							
Molluscum contagiosum virus (MOCV)	421	MOCVgp107	127338-128603	NC_001731	NP_044058		
Genus: Yatapoxvirus							
Tanapox virus (TANV-COD)	152	95L	90056-90514	EF420157	ABQ43725		
Yaba-like disease virus (YLDV-Davis)	152	Ydvgp096	90090-90548	NC_002642	NP_073480		
Yaba Monkey tumor virus (YMTV-Amano)	148	95L	84500-84946	AY386371	AAR07451		
Genus: Cervidpoxvirus							
Deerpox virus (DPV-W848)	151	DpV83gp104	97758-98213	NC_006966	YP_227479		
Genus: Crocodylidpoxvirus	90	CRV125	148670-148942				
Crocodilepox virus (CRV)				NC_008030	YP_784315		

Results

Multiple sequence analysis of CaPV‑A4L gene

PCR amplifcation of the full-length A4L gene of CaPV yielded an amplicon of ~ 500 bp containing an ORF of 486 bp in all of the isolates used in the study. The identity of all PCR products was confrmed by sequencing. A multiple sequence alignment revealed signifcant nt and aa sequence similarity among the Indian and foreign CaPV isolates under study. The nt and aa identity values are listed in Table [3.](#page-4-0) One synonymous mutation and eight non-synonymous mutations were found in the GTPV isolates, and five synonymous mutations and 13 non-synonymous mutations were found in the SSPV isolates. All of the nt and aa changes, including transverse/transition and synonymous/non-synonymous

Table 3 Percentage identity of CaPV A4L coding sequences of Indian GTPV and SPPV isolates to other CaPV isolates

Table 4 List of synonymous and non-synonymous mutations present in CaPV A4L of GTPV and SPPV isolates in this study

mutations are listed in Table [4.](#page-4-1) The presence of guanine (G) at positions 136 and 146, adenine (A) at positions 93 and 216 and cytosine (C) at position 389 may be considered a 'GTPV signature'. Similarly, the presence of A at positions 63, 168, and 276 and G at position 47 can be considered an 'SPPV signature'. The presence of G at position 48 and C at position 98 can be considered an 'LSDV signature'. All SPPV and LSDV isolates possess the 'A' nucleotide substitution at positions 400 and 431, with the exception of SPPV/ Ahemdabad/2009/P5. In contrast to the above, foreign SPPV isolates and LSDV isolates have a G nucleotide substitution at position 465 that is not found in Indian SPPV isolates. A multiple sequence alignment of CaPV ORF095 nucleotide sequences showing specifc signature residues for each of the CaPV members is shown in Fig. [1.](#page-5-0)

Phylogenetic analysis of the CaPV‑A4L gene

A phylogenetic tree based on nucleotide sequences of the CaPV A4L gene is shown in Fig. [2](#page-8-0). The consensus tree

Nucleotide position	47	48	63	93	98	136	146	168	216	273	315	389	400	431	450
Consensus	A	A	\overline{G}	C	A	A	A	G	\overline{c}	G	G	T	A	A	T
Strain															
GTPV/871/Pune/09 (KX398488)				A		G	G		A			С	G	G	
GTPV AV41 (MH381810)				A		G	G		A		Ā	C	G	G	
GTPV/VBRI/2008/P4 (KX398486)				A		G	G		A		A	C	G	G	
GTPV/Sambalpur/2001/P6 (KX398485)				A		G	G		A		A	C	G	G	
GTPV/Avikanagar/P4 (KX398483)				A		$\bf G$	G		A		A	C	G	G	
GTPV 143/Mukteswar/2012 (KX398481)				A		$\overline{\mathbf{G}}$	G		A			C	G	G	
GTPV/Akola/2008/P4 (KX398479)				A		G	G		A		A	C	G	G	
GTPV/Mukteswar/1946/P50 (KX398470)				A		G	G		A			C	G	G	
GTPV/Mukteswar/1946/P5 (KX398469)				A		G	G		A		A	C	G	G	
GTPV FZ (KC951854)				A		$\mathbf G$	G		Â		A	C	G	G	
GTPV/28-08/Palampur/2012 (KX398487)				A		G	G		A		A	C	G	G	
GTPV/Ladakh/2001/P3 (KX398484)				A		Ġ	G		Â			C	G	G	
GTPV 90/Mukteswar/2012 (KX398480)				A		$\mathbf G$	G		A		A	C	G	G	
GTPV/Uttarkashi/1978/P60 (KX398472)				A		G	G		A			C	G	G	
GTPV/Uttarkashi/1978/P101 (KX398471)				A		G	G		A		A	C	G	G	
GTPV KT/Mukteswar/2012 (KX398482)				A		$\bf G$	G		A			C	G	G	
GTPV Gorgan (KX576657)				A		$\mathbf G$	G		A		A	C	G	G	
GTPV G20-LKV (AY077836)				A		$\mathbf G$	G		A		Ă	C	G	G	
GTPV Pellor (AY077835)				A		$\overline{\mathbf{G}}$	$\bf G$		\mathbf{A}		A	$\mathbf C$	G	G	
SPPV/Ahmedabad/2009/P5 (KX398490)	G		$\overline{\mathbf{A}}$					A		$\overline{\mathbf{A}}$		τē	Ġ	Ġ	
SPPV/30-01/Ahmedabad/2008 (KX398494)	$\mathbf G$		A					A		A					A
SPPV/30-02/Ahmedabad/2008 (KX398491)	G	÷	A					A		A					A
SPPV/3174/Pune/2007 (KX398493)	G		A					A		A					А
SPPV/2823/Pune/2007 (KX398492)	G		A					A		A					А
SPPV/Pune/2008/P4 (KX398478)	G		A					A		A					А
SPPV/Ranipet/P54 (KX398474)	G		A					A		A					A
SPPV/32-01/Jalandhar/2011 (KX398495)	G		A					A		A					A
SPPV/Makhdoom/2007/P6 (KX398473)	G		A					A		A					A
SPPV NISKHI (AY077834)	G		A					A		A					А
SPPV A (AY077833)	G		A					A		A					A
SPPV/Srinagar/2000/P5 (KX398476)	G		A					A		A					A
SPPV/Jaipur/India/P3 (KX398489)	G	$\ddot{}$	A					A		A					A
SPPV/Srinagar/2000/P40 (KX398477)	G		A					A		A					А
SPPV/Roumanian Fanar/P37 (KX398475)	G		A					A		A					A
SPPV TU-V02127 (AY077832)	G		A					A		A					A
SPPV-GL (KT438551)	$\bf G$		A					A		A					А
SPPV-GH KT438550)	$\mathbf G$		A												А
LSDV Evros/GR/15 (KY829023)		$\overline{\mathbf{G}}$			$\overline{\mathbf{C}}$										
LSDV SERBIA/Bujanovac/2016 (KY702007)		$\overline{\mathbf{G}}$			Ċ										
LSDV NI-2490 (AF325528)		$\bf G$			Ċ										
LSDV KSGP0240 (KX683219)		G			Ċ										
LSDV NW-LW (AF409137)		$\bf G$			Ċ										
LSDV Cro2016 (MG972412)		G			C				T						
LSDV Neethling-vaccine-OBP (KX764645)		$\bf G$			C				T						
LSDV SIS-Lumpyvax (KX764643)		G			\overline{c}				T						
LSDV Neethling-vaccine LW1959(AF409138)		$\mathbf G$			Ċ				T						

Fig. 1 Multiple sequence alignment of the CaPV A4L gene showing species-specifc signature residues for each member of the genus CaPV indicated by colored boxes

revealed the presence of three distinct genetic clusters separating the three members of the genus *Capripoxvirus* namely GTPV, SPPV and LSDV. This analysis demonstrated that of GTPV and SPPV are more closely related to each other than to LSDV. The signifcance of this grouping/cluster was supported by 1000 replicates of bootstrap testing. All of the Indian GTPV strains except GTPV KT/Mukteswar/2012 clustered together with the Chinese GTPV isolate FZ. GTPV KT/Mukteswar/2012 clustered separately with the foreign GTPV isolates GTPVG20-LKV and GTPV Pellor. The SPPV group showed more intra-cluster diversity than the GTPV and LSDV groups. Unexpectedly, SPPV/Ahmedabad/2009/ P5 formed a diferent subcluster from the other SPPV isolates, although the SPPV isolates in general showed signifcant intra-species sequence similarity. A phylogenetic tree based on the sequences of CaPV-A4L homologs from various animal poxviruses (Table [2\)](#page-3-0) retrieved from the GenBank database is shown in Fig. [3.](#page-9-2) This analysis showed genus-specifc grouping of all chordopoxviruses, with fowlpox virus (FWPV) clustering separately from all other chordopoxviruses. The CaPVs were found to be more closely related to suipoxvirus (SWPV) than to any of the other chordopoxviruses.

Predicted protein characteristics and structural features of CaPV‑A4L

The full-length CaPV-A4L protein (161 aa, 18.68 kDa) was predicted to contain a total of 34 basic, 30 acidic, 47 polar,

and 40 hydrophobic residues and no cysteine residues and to have an isoelectric point of 8.994. A schematic representation of the full length CaPV-A4L is shown in Fig. [4](#page-10-0)A. The predicted antigenic index, hydrophobicity, and surface exposure of amino acid residues obtained using Protean are shown in Fig. [4B](#page-10-0). A secondary structure prediction using PSIRED revealed the presence of six α -helices (α 1, 2-16; aa α2, 22-24; aa α3, 47-62; α4, aa 93-94; α5, aa 100-115; α6, aa 121-150) and only β sheet (aa 71-72) and many coils, as shown in Fig. [4C](#page-10-0). The location of coiled-coil motifs (CCMs) in CaPV-A4L, predicted using COILS2, is shown in Fig. [5A](#page-10-1). A coiled-coil probability plot using scanning windows of 14, 21, or 28 aa residues showed that the central portion of CaPV-A4L can potentially form coiled coils, with aa 54-89 (CCM-I) and aa 100-134 (CCM-II) predicted to have the highest probability using all three window sizes. A 3D structural model was made in I-TASSER after identifcation of templates, and threading alignments. The fnal model had a C-score of -2.14 with an estimated TM score of 0.46 ± 0.15 , and an RMSD score of 9.6 ± 4.6 Å (Fig. [5](#page-10-1)B).

Discussion

Capripoxvirus infection is generally considered to be host-specific in nature [\[6](#page-11-3)], and some strains have been reported to infect sheep, goats and cattle, with more severe disease evident in the homologous host [[7](#page-11-4), [13](#page-11-10), [21\]](#page-11-20). Genus level classifcation based on the host from which the virus was isolated is not reliable due to cross-species transmission, which has been reported frequently in the last decade [\[7,](#page-11-4) [22](#page-11-21)]. The antigenic relatedness of SPPV and GTPV makes them indistinguishable at the serological level, but they can be distinguished using molecular tools, which can help in understanding the molecular epidemiology of these viruses [\[9](#page-11-6), [10,](#page-11-7) [23](#page-11-22), [24\]](#page-11-23). A growing numbers of mixed-fock infections and CaPV-like disease [\[25](#page-11-24)] have worsened the situation, substantiating the need for identifying the causal agent. In this study, ORF095, which encodes a VACV A4L homolog [[12\]](#page-11-9) and is one of the major core immunogenic and conserved proteins of the CaPV genome, was targeted to genetically characterize the viral isolates at the nt and aa level.

Sequence analysis of CaPV-A4L protein revealed that all Indian GTPV isolates have high sequence similarity at both the nt and aa level to GTPV FZ, a Chinese isolate. The same pattern of clustering of Indian GTPVs with Chinese isolates was reported earlier [[26\]](#page-11-25). However, GTPV KT-Mukteswar-2012 contained amino acid substitutions, namely N28S, Y83C and K96R, that were also present in the G20- LKV, Pellor and Gorgan strains. These isolates were identical to each other in the sequenced region at both the nt and aa level. All SPPV isolates from outside of India, namely, the SPPV 10700-99, SPPV A and SPPV Niskhi strains, had a common G substitution at nt position 405 that was not found in any Indian SPPV isolates. Although the AL gene sequence is generally highly conserved among CaPVs, species-specifc signature residues were found that can be used for genotyping or diferentiation of strains within the genus. Being located at the core of CaPV genome, A4L homolog is highly conserved and may not face any immune pressure from the host species that would force genetic changes.

The phylogenetic analysis showed that CaPVs formed three distinct clusters, namely GTPV, SPPV and LSDV, as observed for other genes used for the characterization of CaPVs [[9,](#page-11-6) [27\]](#page-11-19). The topology of the consensus tree also revealed that GTPV and SPPV are more closely related to each other than to LSDV, supporting the earlier conclusions based on analysis of whole genome sequences of CaPVs [[11\]](#page-11-8). It has been suggested recently that GTPV and LSDV are more closely related to each other than to SPPV and that they emerged from a common ancestor closely related to SPPV. This was based on phylogenetic studies on diferent genomic segments [\[9](#page-11-6), [11,](#page-11-8) [28](#page-11-26)]. More studies are needed to clarify the evolutionary history of CaPV.

CaPV-A4L is a highly conserved immunodominant core protein. Therefore, it is not surprising that the nt sequence of its gene and its deduced aa sequence were found to be more than 96.3% and 92.0% identical, respectively, in all CaPVs. Multiple alignments of A4L sequences and phylogenetic analysis did not reveal specifc diferences among the vaccine and virulent strains of GTPV and SPPV, indicating that the A4L might not play a role in virulence or in the attenuation process. Subdivisions or subclusters were found in the GTPV, SPPV and LSDV groups with considerable bootstrap support, which is in agreement with previously reported studies examining the GPCR and RPO30 genes [[23,](#page-11-22) [28](#page-11-26)]. One interesting feature of the A4L gene phylogeny is that the SPPV/Ahmedabad/2009/P5 isolate forms a separate subcluster between the GTPV and SPPV clusters. Multiple sequence alignment also revealed that this particular SPPV isolate had four nucleotide changes (T389C, A400G, A431G and A450T) in common with GTPV isolates. Therefore, sequencing of the entire genome of this isolate is warranted to investigate whether recombination might have occurred between SPPV and GTPV. Also, the C216T mutation was only observed in the LSDV Cro/2016, OBP, SIS Lumpyvax and NW-LW1959 isolates, which form a separate subcluster from other LSDV isolates in the phylogenetic tree (Fig. [1](#page-5-0))

The ORF095 protein is a homolog of myxoma virus (MYXV) M093L [[29](#page-11-27)] and vaccinia virus (VACV) A4L [[13\]](#page-11-10). These viruses encode a 39-kDa acidic virion core protein that is synthesized at a late stage after infection. The late viral protein of VACV, designated as p39, corresponds to the product of A4L in VACV Copenhagen and A5L in VACV WR strains and was reported previously to induce a strong persistent immune response in humans [\[30](#page-11-28), [31](#page-11-29)]. The

Fig. 2 Phylogenetic tree based on nucleotide sequences of CaPV-A4L ◂genes of Indian GTPV and SPPV isolates and other CaPV sequences available in the GenBank database (Table [1](#page-2-0)). The tree was constructed by the neighbor-joining method in MEGA version 7.0 with 1000 bootstrap replicates. Symbols in the fgure indicate vaccine strains of GTPV and SPPV used in India

co-localization of A5/A4 with replicated DNA in virosomes and its co-purifcation with nuclease activity indicate a possible role of this core protein in DNA processing or packaging [[13\]](#page-11-10). Also, the A4L or p39 protein does not undergo post-translational modifcations and acts as a matrix-like linker between the core and the innermost of the two membranes of the intracellular mature VACV virion [\[31\]](#page-11-29) and makes a part of the spike-like protrusions present on the IMV core. Similarly, CaPV-A4L is located in the core but is partially exposed to the host immune system, making it potentially useful for diferentiation of CaPV-vaccinated and infected hosts. This protein should be overexpressed in a suitable heterologous host and evaluated in an immunoassay such as ELISA as a coating antigen.

The CaPV-A4L protein was predicted to have a high antigenic index with high hydrophilicity and surface probability, suggesting its potential use for diagnostic or prophylactic purposes. The CaPV-A4L homolog protein of GTPV Uttartkashi (vaccine strain) has 93.79 to 100% sequence identity among members of the genus *Capripoxvirus*, with the highest identity observed with the GTPV FZ strain and the lowest with LSDV isolates. A4L homologs from other poxviruses had shown a low percentage of identity, viz. swinepox virus (50.31%), deerpox virus (44.87-45.51%), Yaba-like and tanapox viruses (37.65%), vaccinia-like viruses (37.11-37.18%), followed by myxoma and rabbit fbroma viruses (37.11%) with least identity to fowlpox virus (FWPV). FWPV strains carry a VACV A4L homolog, which is an immunodominant, 39K core protein possessing highly charged domains at each end of the protein and a

12-amino-acid serine-rich repeat sequence in the middle of the protein [\[32](#page-11-30)]. This repeat structure has also been reported to be present in a molluscum contagiosum virus homolog (MC107L) but absent from VACV-A4L. As expected, CaPV-A4L did not show any such repeat sequence in its primary structure. The complete A4L protein may be essential for the virus, as it plays a major role in virus assembly. This fact is corroborated by evidence that shRNA targeting CaPV A4L can efectively inhibit the replication of GTPV in Vero cells, resulting in a 463.5-fold reduction *in vitro* [[12\]](#page-11-9). This presents a new anti-GTPV strategy for control of goatpox in the future. Strong coiled coil motifs (CCM-I and II) responsible for oligomerization under native conditions are present in the middle region of the protein. Oligomerization might stabilize the structure and function of the protein under common storage conditions [[33\]](#page-11-31). This might be one of the potential attributes of a candidate diagnostic or prophylactic antigen [[34,](#page-11-32) [35\]](#page-11-33).

Conclusions

In this study, we found that the CaPV-A4L homolog is highly conserved among CaPVs but shows genetic variations that can be can be exploited for genotyping by phylogenetic analysis and testing for signature residues in the gene sequence. Comparative sequence analysis of A4L homologs of diferent pox viruses revealed a low percentage of overall identity but highly conservation at the genus level. The oligomerization ability of this protein predicted in this study needs to be tested experimentally. Potential signature residues specifc for each CaPV member are indicative of an alternate genotyping strategy. Also, the ideal protein characteristics revealed by A4L structural analysis show that it

Fig. 3 Phylogenetic tree of poxviruses based on the deduced amino acid sequence of the A4L protein. Representative poxviruses and their A4L amino acid sequences (Table [2](#page-3-0)) were used to construct

can be a potential candidate antigen for heterologous expression and application for immunodiagnostic or prophylactic purposes.

Acknowledgements The authors thank the Director, Indian Veterinary Research Institute, for providing necessary facilities to carry out this work, and the staff of the pox virus laboratory, IVRI, Mukteswar, for their valuable and timely help in carrying out this work. The authors also thank the Directors and staff of the regional disease diagnostic laboratories of the state animal husbandry departments for sending the clinical samples to the pox laboratory, IVRI, Mukteswar, for diagnostic investigation of animal viral diseases. This study is part of the MVSc thesis work of the frst author. The fnancial support provided by Indian Veterinary Research Institute, Izatnagar, under institute project (IXX09899) is also acknowledged.

the tree by the neighbour-joining method in MEGA version 7.0 with 1000 bootstrap replicates

Compliance with ethical standards

Conflict of interest The authors declare that there is no confict of interest.

References

- 1. International Committee on Taxonomy of Viruses (ICTV) [https](https://talk.ictvonline.org/taxonomy/) [://talk.ictvonline.org/taxonomy/](https://talk.ictvonline.org/taxonomy/)
- 2. Office Internationale des epizooties (World Health Organization for animals) WAHIS portal: animal health data. [https://www.](https://www.oie.int/en/animal-health-in-the-world/wahis-portal-animal-health-data) [oie.int/en/animal-health-in-the-world/wahis-portal-animal-healt](https://www.oie.int/en/animal-health-in-the-world/wahis-portal-animal-health-data) [h-data](https://www.oie.int/en/animal-health-in-the-world/wahis-portal-animal-health-data)

Fig. 4 Predicted characteristics of the virion core protein CaPV-A4L. **a** Schematic diagram of the aPV ORF095 protein sequence, which is 161 aa residues (18.6 kDa) in length; **b** Predicted protein characteristics. (a) Kyte-Doolittle hydro philicity plot, (b) Jameson-Wolf antigenic index, (c) Emini sur face probability plot. **c** GTPV-A4L sequence with predicted secondary structure. The sec ondary structure was predicted using PSIPRED. Pink tubes indicate α-helices. β-strands are represented by yellow arrows. Sec str, secondary structure; aa $#$, amino acid number; α , helix region; β, strand region. The helix and strand regions are numbered sequentially

Fig. 5 Structural prediction of the CaPV A4L protein. **a** Probability plot for CCMs in GTPV-A4L predicted using COILS2 program. Peaks in the graph indicate regions of higher coiled-coil probability. The default output of prob abilities with scanning windows of 14, 21 and 28 aa residues is shown in green, blue, and red, respectively. **b** Predicted GTPV-A4L protein model made using the SPICKER program of I-TASSER. The closest structural analogs were myosin/ tropomyosin contractile proteins (pdb hits: 2tmaB and 2dfsA)
with TM alignment scores of > 0.630 and 0.621, respectively

² Springer

- 3. Madhavan A, Venkatesan G, Kumar A (2016) Capripoxviruses of small ruminants: current updates and future perspectives. Asian J Anim Vet Adv 11:757–770
- 4. Garner MG, Sawarkar SD, Brett EK, Edwards JR, Kulkarni VB, Boyle DB, Singh SN (2000) The extent and economic impact of sheep pox and goat pox in the state of Maharashtra, India. Trop Anim Health Prod 32:205–223
- 5. Yeruham I, Yadin H, Van Ham M, Bumbarov V, Soham A, Perl S (2007) Economic and epidemiological aspects of an outbreak of sheeppox in a dairy sheep fock. Vet Rec 160(7):236–237
- 6. Babiuk S, Bowden TR, Boyle DB, Wallace DB, Kitching RP (2008) Capripoxviruses: an emerging worldwide threat to sheep, goats and cattle. Transbound Emerg Dis 55:263–272
- 7. Bhanuprakash V, Venkatesan G, Balamurugan V, Hosamani M, Yogisharadhya R, Chauhan RS, Pande A, Mondal B, Singh RK (2010) Pox outbreaks in sheep and goats at Makhdoom (Uttar Pradesh), India: evidence of sheep pox virus infection in goats. Transbound Emerg Dis 57:375–382
- 8. Bhanuprakash V, Hosamani M, Singh RK (2011) Prospects of control and eradication of capripox from the Indian subcontinent: a perspective. Antivir Res 91:225–232
- 9. Hosamani M, Mondal B, Tembhurne PA, Bandyopadhyay SK, Singh RK, Rasool TJ (2004) Diferentiation of sheep pox and goat poxviruses by sequence analysis and PCR-RFLP of P32 gene. Virus Genes 29:73–80
- 10. Venkatesan G, Balamurugan V, Yogisharadhya R, Kumar A, Bhanuprakash V (2012) Diferentiation of sheeppox and goatpox viruses by polymerase Chain reaction-restriction fragment length polymorphism. Virol Sin 27(6):353–359
- 11. Tulman ER, Afonso CL, Lu Z, Zsak L, Sur JH, Sandybaev NT, Kerembekova UZ, Zaitsev VL, Kutish GF, Rock DL (2002) The genomes of sheeppox and goatpox viruses. J Virol 76(12):6054–6061
- 12. Zhao Z, Wu G, Zhu X, Yan X, Dou Y, Li J, Zhu H, Zhang Q, Cai X (2012) RNA interference targeting virion core protein ORF095 inhibits goatpox virus replication in Vero cells. Virol J 9:48
- 13. Maa JS, Esteban M (1987) Structural and functional studies of a 39,000-*M*r immunodominant protein of vaccinia virus. J Virol 61:3910–3919
- 14. Bowden TR, Coupar BE, Babiuk SL, White JR, Boyd V, Duch CJ, Shiell BJ, Ueda N, Parkyn GR, Copps JS, Boyle DB (2009) Detection of antibodies specifc for sheeppox and goatpox viruses using recombinant capripoxvirus antigens in an indirect enzyme-linked immunosorbent assay. J Virol Methods 161(1):19–29
- 15. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specifc gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
- 16. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874
- 17. Jones DT (1999) PSIPRED: protein secondary structure prediction based on position-specifc scoring matrices. J Mol Biol 292(2):195–202
- 18. Lupas A, Van Dyke M, Stock J (1991) Predicting coiled coils from protein sequences. Science 252:1162–1164
- 19. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinform 9:40
- 20. Yang J, Zhang Y (2015) I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res 43:W174–W181
- 21. Tuppurainen ES, Pearson CR, Bachanek-Bankowska K, Knowles NJ, Amareen S, Frost L, Henstock MR, Lamien CE, Diallo A, Mertens PP (2014) Characterization of sheep pox virus vaccine for cattle against lumpy skin disease virus. Antivir Res 109:1–6
- 22. Santhamani R, Venkatesan G, Minhas SK, Shivachandra SB, Muthuchelvan D, Pandey AB, Ramakrishnan MA (2015) Detection and characterization of atypical capripoxviruses among small ruminants in India. Virus Genes 51(1):33–38
- 23. Lamien CE, Le Goff C, Silber R, Wallace DB, Gulyaz V, Tuppurainen E, Madani H, Caufour P, Adam T, El Harrak M, Luckins AG, Albina E, Diallo A (2011) Use of the Capripoxvirus homologue of Vaccinia virus 30 kDa RNA polymerase subunit (RPO30) gene as a novel diagnostic and genotyping target: development of a classical PCR method to diferentiate Goatpoxvirus from Sheep poxvirus. Vet Microbiol 149:30–39
- 24. Venkatesan G, Balamurugan V, Bhanuprakash V (2014) Multiplex PCR for simultaneous detection and differentiation of sheeppox, goatpox and orf viruses from clinical samples of sheep and goats. J Virol Methods 195:1–8
- 25. Greth A, Gourreau JM, Vassart M, Nguyen-Ba-Vy, Wyers M, Lefevre PC (1992) Capripoxvirus disease in an Arabian oryx (*Oryx leucoryx*) from Saudi Arabia. J Wildl Dis 28(2):295–300
- 26. Zeng X, Chi X, Li W, Hao W, Li M, Huang X, Huang Y, Rock DL, Luo S, Wang S (2014) Complete genome sequence analysis of goatpox virus isolated from China shows high variation. Vet Microbiol 173(1–2):38–49
- 27. Karki M, Kumar A, Venkatesan G, Arya S, Pandey AB (2018) Genetic analysis of L1R myristoylated protein of Capripoxviruses reveals structural homogeneity among poxviruses. Infect Genet Evol 58:224–231
- 28. Le Goff C, Lamien CE, Fakhfakh E, Chadeyras A, Aba-Adulugba E, Libeau G, Tuppurainen E, Wallace DB, Adam T, Silber R, Gulyaz V, Madani H, Caufour P, Hammami S, Diallo A, Albina E (2009) Capripoxvirus G-protein-coupled chemokine receptor: a host-range gene suitable for virus animal origin discrimination. J Gen Virol 90(8):1967–1977
- 29. Van Vliet K, Mohamed MR, Zhang L, Villa NY, Werden SJ, Liu J, McFadden G (2009) Poxvirus proteomics and virus-host protein interactions. Microbiol Mol Biol Rev 73:730–749
- 30. Demkowicz WE, Maa JS, Esteban M (1992) Identifcation and characterization of vaccinia genes encoding proteins that are highly antigenic in animals and are immunodominant in vaccinated humans. J Virol 66:386–398
- 31. Cudmore S, Blasco R, Vincentelli R, Esteban M, Sodeik B, Griffthsm G, Locker JK (1996) A vaccinia virus core protein, p39, is membrane associated. J Virol 70:6909–6921
- 32. Boulanger D, Green P, Smith T, Czerny CP, Skinner MA (1998) The 131-amino-acid repeat region of the essential 39-kilodalton core protein of fowlpox virus FP9, equivalent to vaccinia virus A4L protein, is nonessential and highly immunogenic. J Virol 72(1):170–179
- 33. Chang T, Chang S, Hsieh F, Ko T, Lin C, Ho MR, Wang I, Hsu ST, Guo RT, Chang W, Wang AH (2013) Crystal structure of vaccinia viral A27 protein reveals a novel structure critical for its function and complex formation with A26 protein. PLoS Pathog 9:e1003563
- 34. Kumar A, Yogisharadhya R, Bhanuprakash V, Venkatesan G, Shivachandra SB (2015) Structural analysis and immunogenicity of recombinant major envelope protein (rA27L) of bufalpox virus, a zoonotic Indian vaccinia-like virus. Vaccine 33(41):5396–5405
- 35. Dashprakash M, Venkatesan G, Kumar A, Sankar M, Arya S, Ramakrishnan MA, Pandey AB, Mondal B (2019) Prokaryotic expression, purifcation and evaluation of goatpox virus ORF117 protein as a diagnostic antigen in indirect ELISA to detect goatpox. Arch Virol 164(4):1049–1058
- 36. Mathijs E, Vandenbussche F, Haegeman A, Al-Majali A, De Clercq K, Van Borm S (2016) Complete genome sequence of the goatpox virus strain Gorgan obtained directly from a commercial live attenuated vaccine. Genome Announc 4(5)
- 37. Kara PD, Afonso CL, Wallace DB, Kutish GF, Abolnik C, Lu Z, Vreede FT, Taljaard LC, Zsak A, Viljoen GJ, Rock DL (2003) Comparative sequence analysis of the South African vaccine strain and two virulent feld isolates of Lumpy skin disease virus. Arch Virol 148(7):1335–1356
- 38. Tulman ER, Afonso CL, Lu Z, Zsak L, Kutish GF, Rock DL (2001) Genome of lumpy skin disease virus. J Virol 75:7122–7130
- 39. Mathijs E, Vandenbussche F, Haegeman A, King A, Nthangeni B, Potgieter C, Maartens L, Van Borm S, De Clercq K (2016) Complete genome sequences of the Neethling-like lumpy skin disease virus strains obtained directly from three commercial live attenuated vaccines. Genome Announc 4(6):e01255–16
- 40. Agianniotaki EI, Mathijs E, Vandenbussche F, Tasioudi KE, Haegeman A, Iliadou P, De Clercq K (2017) Complete genome sequence of the lumpy skin disease virus isolated from the frst reported case in Greece in 2015. Genome Announc 5(29)
- 41. Vandenbussche F, Mathijs E, Haegeman A, Al-Majali A, Van Borm S, De Clercq K (2016) Complete genome sequence of

Capripoxvirus strain KSGP 0240 from a commercial live attenuated vaccine. Genome Announc 4(5):e01114–e01116

- 42. Toplak I, Petrovic T, Vidanovic D, Lazic S, Sekler M, Manic M, Kuhar U (2017) Complete genome sequence of lumpy skin disease virus isolate SERBIA/Bujanovac/2016, Detected during an outbreak in the Balkan Area. Genome Announc 5(35)
- 43. Lojkic I, Simic I, Kresic N, Bedekovic T (2018) Complete genome sequence of a lumpy skin disease virus strain isolated from the skin of a vaccinated animal. Genome Announc **6**(22).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.