Virus Genes (2006) 32:71–78
DOI 10.1007/s11262-005-5847-y

ORIGINAL ARTICLE

Sequence analysis of the haemagglutinin and fusion protein genes of peste-des-petits ruminants vaccine virus of Indian origin*

P. Dhar · D. Muthuchelvan · A. Sanyal · R. Kaul · R. P. Singh · R. K. Singh · S. K. Bandyopadhyay

Received: 4 April 2005 / Revised and Accepted: 27 July 2005

Abstract he amino acid composition of the two surface proteins of peste-des-petits ruminants vaccine virus belonging to lineage four from India were deduced from the nucleotide sequence. The fusion (F) protein gene of PPRV Sungri/96 is 2405 nucleotides long and in relation to the length, it is 80 nucleotides longer than that of PPRV Nigeria/75/1 which are found to be present at the 5′UTR of this virus. The complete F gene alignment with other morbillivirus reveals a homology of 89% with PPRV/Nigeria/75/1 and 48–51% with other morbilliviruses. The F protein of PPRV Sungri/96 exhibited characteristics similarity to those of other morbillivirus F proteins. The overall amino acid similarity with its counterpart PPRV Nigeria/75/1 was 96%; with other morbilliviruses it is 65–74%. The PPRV Sungri/96 haemagglutinin (H) protein gene is 1954 nucleotides long and showed a sequence homology of 90.7% with PPRV/Nigeria/75/1 and with other morbilliviruses it ranged from 33% to 45%. At amino acids level, PPRV Sungri/96 showed a homology of 92.3% with PPRV/Nigeria/75/1 and 34–49% with other morbilliviruses. The phylogenetic tree constructed for F and H gene reveals four separate groups which is very similar to that found in other genes. To the best of our knowledge this is the first report describing the F and H genes of an Indian isolate.

Keywords Peste-des-petits ruminants virus · F gene · H gene · Nucleotide sequence · Amino acid sequence · Morbillivirus · Vaccine virus

*The sequence data reported in this paper have been submitted to the GenBank and have been assigned the Accession Number: AY560591. P. Dhar and D. Muthuchelvan equally contributed to the work.

P. Dhar
Indian Veterinary Research Institute, Izatanagar, 243 122, Bareilly, India

D. Muthuchelvan
Central Institute of Fisheries Technology, 682 029, Cochin, India

A. Sanyal
Project Directorate on Foot-and-Mouth Disease, IVRI Campus Mukteswar-Kumaon, 263 138, Nainital, India

R. Kaul
Indian Veterinary Research Institute, Mukteswar Campus, 263 138, Nainital, India

R. P. Singh
Indian Veterinary Research Institute, Izatanagar, 243 122, Bareilly, India

R. K. Singh
Indian Veterinary Research Institute, Mukteswar Campus, 263 138, Nainital, India

S. K. Bandyopadhyay (✉)
Animal Husbandry Commission, 234A, Krishi Bhavan, 110 001, New Delhi, India
e-mail: skbandyahc@rediffmail.com

Introduction

Peste des petits ruminants (PPR) is an acute and febrile viral disease of small ruminants, which is characterized by pyrexia, ocular and nasal discharges, erosive stomatitis and diarrhoea [1]. It was first reported in India in 1987 [2] and became endemic in the country [3]. India being a vast country with a population of more than 130 million goats and 58 million sheep; the PPR is considered as a main constraint in augmenting the productivity of small ruminants in this nation [3].

The causative agent, *peste des petits ruminants* virus (PPRV) is classified as a member of the genus Morbillivirus, Family *Paramyxoviridae* under the order Mononegavirales [4]. The other members of this genus include rinderpest virus (RPV), measles virus (MV), canine distemper virus (CDV), phocine distemper virus (PDV), dolphin morbillivirus (DMV) and porpoise morbillivirus (PMV) [5, 6]. The genome of morbilliviruses is a single stranded-RNA, ~16 kb long with negative polarity [7]. It is divided into six transcriptional units encoding two non-structural proteins (V, C) and six structural proteins: the surface glycoproteins (F and H), the matrix protein (M), the nucleoprotein (N), the phosphoprotein (P) which forms the polymerase complex in association with the large protein (L) [8, 9]. Presently, the genome sequences of most of the morbilliviruses are known; however, in the case of PPRV the nucleotide sequence of only the African vaccine virus Nigeria75/1 is nearly completed [1, 6, 8, 10, 11].

Based on the molecular epidemiological studies PPRV isolates have been classified in to four lineages (Lineage 1, 2, 3 & 4). Of which lineages 1, 2 & 3 is found in Africa and the lineage 4 is exclusively reported to occur in the Middle East, Arabia-and Indian sub-continent. The vaccine virus Nigeria75/1 (named here as PPRV-N) belongs to lineage 1 [12, 13]. The ''PPRV Sungri/96'' (named here as PPRV-S) virus used in the present study belongs to lineage 4 [12] and has been extensively characterized using a panel of monoclonal antibodies directed against different proteins [14] and thermostability [15]. Recently, a homologous live-attenuated vaccine has been developed at Indian Veterinary Research Institute (IVRI) using PPRV Sungri/96, which has been used in the field for vaccination of sheep and goats in India [I5].

Since not much work has been done to understand the genetic makeup of a lineage 4 virus, our lab is engaged in completing the genome sequencing of the vaccine virus PPRV-S. In this context here we report the fusion and haemagglutinin gene sequence of PPRV-S and its comparative analysis with other morbilliviruses including the lineage 1 PPR vaccine virus PPRV-N of African origin. The vaccine developed at IVRI is likely to be used extensively in field conditions in India and other Asian countries. Genetic characterization may help in better understanding of the vaccine virus and also to assess a vaccine-related outbreak.

Materials and methods

Cells and virus

Vero cells at passage level of 131 were used for propagation of the vaccine seed virus developed at Rinderpest Laboratory; IVRI-Mukteswar campus using an indigenous isolate of PPR virus (''PPR Sungri/96''). The vaccine virus was propagated as described previously [15] and aliquots of 250 μl were stored at −80°C until use.

RNA extraction, cDNA synthesis and PCR

Total RNA was extracted by the acid–guanidium-thiocyanate–phenol–chloroform method essentially as described earlier [12, 16]. The oligonucleotides used in RT-PCR to generate overlapping fragments covering the entire Fusion and Haemagglutinin protein genes are shown in Table 1 and were designed from the PPRV Nigeria75/1 sequences available in the GenBank database (Accession Numbers: Z47977, Z81358, Z37017) using the software DNASIS version 2.6 (Hitachi, Japan) and were obtained from M/S Metabion GmbH, Germany. Reverse transcription was performed on 1–5 μg of total RNA using MMLV reverse transcriptase and random hexamers (75 μg) at 37°C for 1 h and subsequent PCR amplification was carried out using 5 μl of the RT product. The PCR cycling conditions were as described previously [13, 16]. The cDNAs were

Table 1 Primers used for amplification of F and H genes of PPRV-S in the study

Sl.No	Name of the primer	Sequence (5′ to 3′)	T_A	Binding location in the gene	Expected size of the product
1	ppr_mfor2	GATACTCCCCAGAGATTCAG	55°C	480–500 in M gene	1294 bp
	Pprf_rev1	GTGGTCCTCCCTCGGTCTGT	55°C	278–258 in F gene	
2	Pprf_for2	AGAACCCTCCCCCCCGCACA		228–248 in F gene	475 bp
	Pprf_rev4	ATAACTAGAGTTTGGTGGCTTGG	55°C	703–681 in F gene	
3	Pprf_for4	GGCCAAGCCACCAAACTCTA	55°C	679–699 in F gene	935 bp
	pprf_rev10	GAAGCATTCCTGAAGCAATGGG	55°C	1614–1593 in F gene	
4	Pprf_for7	AATATCGAATTTCGATGAGAC	55°C	1512–1531 in F gene	964 bp
	PDPPRFHREV	CAGTAGGGATCCAAGTAAG	55°C	145–127 in H gene	
5	Pprh_13	ACGAAAGGTCAATCACCATG	55°C	1–20 in H gene	1848 bp
	Pprh_re2	GACTGGATTACATGTTACCT	55°C	1848–1829 in H gene	
6	Pprh_fr3	ATTCGGAAGAACATATACTGTCT'GG	55°C	695–719 in H gene	718 bp
	Pprh_re4	CTAGAAATGACTGCTCATAAGGTGG	55°C	1413– 1389 in H gene	
7	Pprh_fr4	TCATCATCTTACTTCTACCCAGTCC	55°C	1663–1687 in H gene	
8	Pprl_rev1	CGTGTCAATGCCACAATGCTC	55°C	994–973 in L gene	

subjected to a 30-cycle amplification (denaturation at 95°C, annealing for 1 min at 50°C and primer extension for 2 min at 72°C) using 10 pmol of each primer.

Cloning of PCR products and sequencing

The PCR amplicons were checked for its correct size in 1% agarose gel and purified using Wizard® PCR Purification system (Promega). The purified amplicons were cloned in to pGEM-T vector and the recombinant plasmid DNA was isolated from representative clones and checked for its correct size as described previously [12]. These cloned amplicons were sequenced on both the strands using fmol DNA cycle sequencing kit (Promega) and cy5-labeled M13 forward and reverse primers in an automated sequencer ALF express II, (Amersham Pharmacia Biotech, U.K.).

Sequence and phylogenetic analysis

The sequenced fragments of the gene were assembled using the Megalign software of the DNASTAR package. The portions overlapping and the primer sequences were eliminated appropriately. For comparison, the following sequences (GenBank Accession numbers in parenthesis) were obtained from NCBI sequence databases; vaccine virus PPRV Nigeria/75/1 (Z81358, Z37017), rinderpest virus (Z30697), Dolphin Morbillivirus (NC005283) phocine distemper virus strain PDV/DK88 (X75717), canine distemper virus (AF305419), and Measles virus (AF266289). The phylogenetic tree was constructed using the Neighbour Joining method with Kimura−2-parameter model available in the program MEGA version 2.1 [17]. The alignment gaps were excluded from pairwise distance estimations. The robustness of the predicted tree was statistically evaluated using the boot trap method [18, 19]. The bootstrap *P*-values are obtained after 10,000 replications. The sequence data reported in this paper have been submitted to the GenBank (Accession Number AY560591).

Results and discussion

Fusion protein gene

The entire F gene of PPRV-S composed of 2405 nucleotides including the poly-(A) tail. The gene starts with the semi-conserved gene start sequence AGGG and ends with AAAC, which is the same for F gene of PPRV-N. Compared to the published nucleotide sequences of the F-protein gene of other morbilliviruses, the PPRV-S F gene is found to be the longest gene reported so far. It is longer than PPRV-N, RPV, MV, CDV/PDV and DMV genes by 80, 52, 24, 200 and 189 nucleotides, respectively [11, 20–25]. The additional 80 nucleotides are present at the 5′ untranslated region (UTR) at position 11–90 of PPRV-S. Similar sequence could be found in the PPRV turkey isolate whose sequence is available in the GenBank database (Accession number AJ849636)

The complete PPRV Sungri F gene alignment with other morbillivirus (Table 2) reveals a homology of 89% with PPRV-N and 48–51% with other morbilliviruses. As with other morbilliviruses, the PPRV-S F gene contained a long stretch of 628 nucleotides rich in G-C residues (68.6%) at the 5′ UTR. The ORF starts at position 629–631 and ends at 2267–2269. At the 5′ UTR, PPRV-S shares low to moderate range of nucleotide sequence homology with other morbilliviruses, and it was 25.6% for DMV and 80.3% for PPRV-N [11]. As the sequence divergence observed in this region is ~20% between PPRV-S and PPRV-N, the usefulness of this region for molecular epidemiology should be explored. The 3′ UTR is of 136 nucleotides long ending at AAACAAAA, which is followed by the intergenic trinucleotide CTT.

The PPRV-S F-protein consists of 546 amino acids similar to that of PPRV-N and RPV F proteins, with a predicted molecular weight (MW) of 59.137 kDa. Comparison of the PPRV fusion protein amino acid sequences as a whole with those of other morbilliviruses (Table 2) showed that the lowest homology was with PDV at 65.6% and the highest homology with PPRV-N at 96.2%.

Table 2 Similarity between the F protein and.H protein of PPRV Sungri/96 and other morbilliviruses

Virus	F gene nucleotide	F ORF nucleotide	F amino acid sequence	F peptide sequence	F2 peptide sequence	H gene nucleotide	H ORF nucleotide	H amino acid sequence
MV	50.3	60.6	72.9	76.5	68.5	44.0	43.4	47.2
RPV	51.8	61.8	74.4	76.3	77.5	45.6	45.4	48.9
CDV	46.2	53.6	66.5	69.6	64.0	33.9	36.7	34.7
PDV	46.8	55.6	65.6	68.9	61.8	35.8	37.6	36.2
DMV	51.2	61.6	74.2	77.4	75.3	42.0	42.4	46.1
PPRV Nigeria	89.7	93.5	96.2	96.8	98.9	90.7	92.0	92.3

The values for nucleotides and amino acid are given as percent similarity/identity

The alignment of the different F-protein sequences (Fig. 1) revealed a high degree of sequence conservation apart from two main domains and nearly identical hydropathic profiles as observed previously for morbilliviruses [11, 22]. The first domain, signal peptide in the N-terminus of the protein consists of 19 amino acids, out of which 5 residues are variable between PPRV isolates. In the second non-conserved long domain (aa 485–517) which includes the hydrophobic anchor membrane sequence (aa 485–502), only two amino acid variations could be observed between

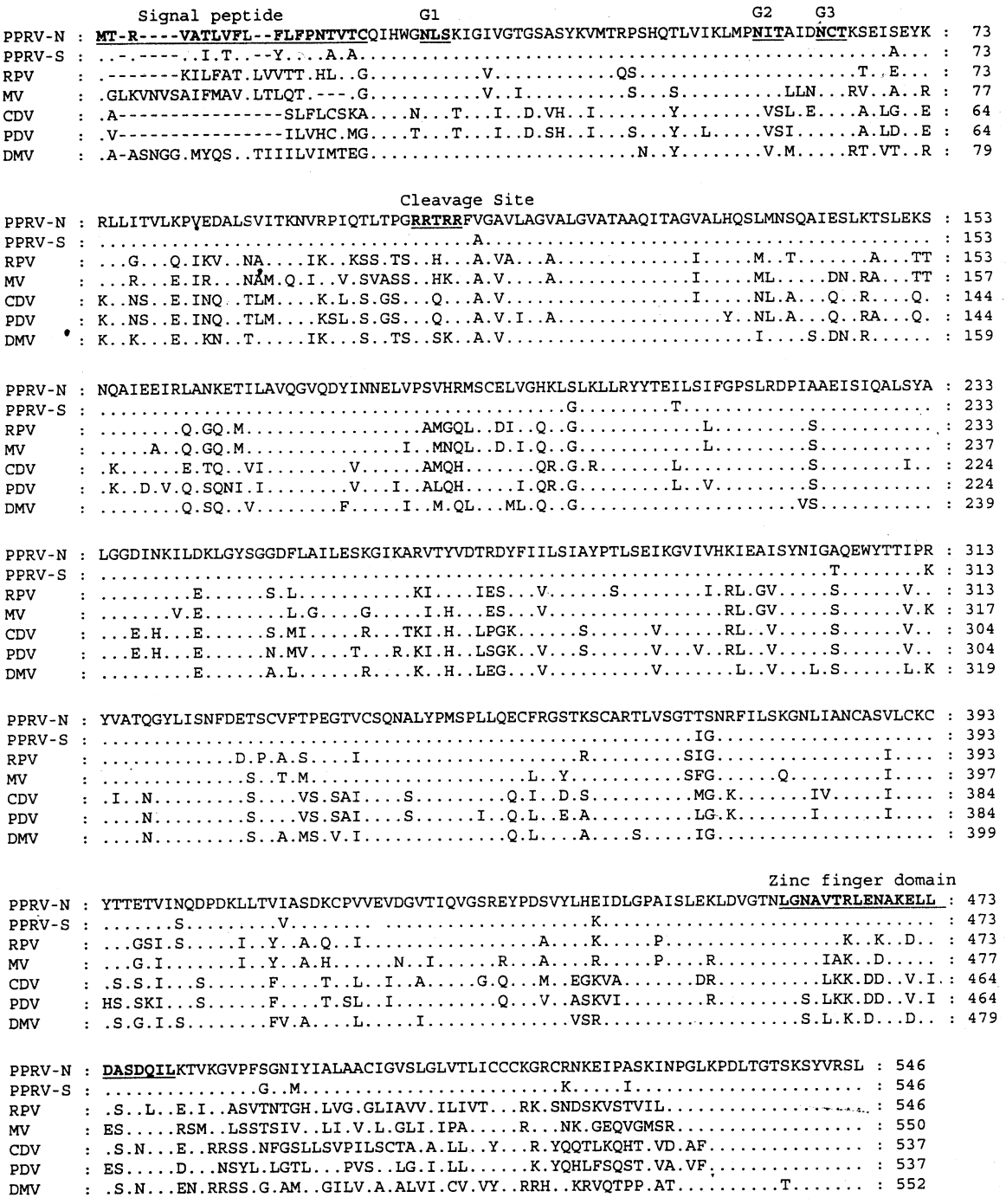

Fig. 1 Comparison of the predicted amino acid (aa) sequence of F protein of PPRV-S with PPRV-N and other morbilliviruses. Dot (.) represent the identity with PPRV-N; dash (-) denotes gap generated during alignment and the difference in aa sequence is represented by a single letter aa code. The signal peptide, potential glycosylation sites (G1–G3), F0 cleavage site and Zinc finger domain are marked appropriately

```
                                                    Membrane anchor domain
PPRV-N : MSAQRERINAFYKDNLH***NKTH***RVILDRERLTIERPY***ILLGVLLVMFLSLIGLLAIAGI***RLHRATVGTAEIQSRLNTNIEL :  80
PPRV-S : .............G.P...N..........V.....................................L..... :  80
RPV    : ..PP.D.VD.Y....FQF.NT..V.NK.Q.L....CM..T..F......V.............A.N..K.NND.T.S.DI :  80
MV     : ..P..D.........P.P.GS.IVIN..H.M.D...V..A..F.....................AIY....HKS.S..LDV :  80
CDV    : .LPYQDKVG......ARANSTKLS.VT.GHGGR..PY..F...ILLVGILA....T.V.F.QVSTSNM.FSRL.KEDM.K :  80
PDV    : .FSHQDKVG....N.ARANSSKLP.VTDEVEER.SPWF.SI..ILLVGILI....T...F.QVVKSNL.FNKL.IEDM.K :  80
DMV    : ..SP.DKVD.....IPRPRNN..L..N..VI....L..V...A......V.......V..QK..TNSI.VNRK.S..L.T :  80

PPRV-N : TESIDHQTKDVLTPLFKIIGDEVGIRIPQKFSDLVKFISDKIKFLNPDREYDFRDLRWCMNPPERVKINFDQFCEYKAAV : 160
PPRV-S : ..................................................................X... : 160
RPV    : .K..EY.V..............L.T..R.T..T.............K......IN..I.....I..DY..Y.AHT..E : 160
MV     : .N..E..V..............L.T..R.T.......................T..I.....I.LDY..Y.ADV..E : 160
CDV    : S.AVH..VI.............I.L.L...LNEIKQ..LQ.TN.F..N..F.....H..I...ST..V..TNY..SIGIR : 160
PDV    : ..AVH..V..............L.L...LNEIKQ..VQ.TN.F..N..F...E.H..I...SK..V..T.Y..ITEFR : 160
DMV    : .V..E.HV.....'.....L.M...LTEIMQ...N.............N..H..V...DQ...DYA.Y.NHI..E : 160

PPRV-N : KSVEHIFESSL***NRSE***RLRLLTLGPGTGCLGRTVTRAQFSELTLTLMDLDLEIKH***NVSS***VFTVVEEGLFGRTYTVWRSDTG : 240
PPRV-S : ..I......P..K.KK.QS.........Q.......H.............M.........................AR : 240
RPV    : DLITMLVN...TGTTVP.TSLVNL.RN.T.P.T.KG...NIS...SGIYSGRGY.I..MI.ITGK.MY.S..L.GKYNQR : 240
MV     : ELMNALVN.T.LETRTTNQFLAVSKGN.S.P.TI.G...NMS.S.L..Y.GRGY....IV.MTSQ.MY.G..L.EKPNLS : 240
CDV    : .AIASAANPI.LSALSGGRGDIF.PHR.S.A.TSVGKVFP.SVS.SMSLISRTSE.INML.AISD.VY.K..LLVPD.IE : 240
PDV    : EATRSVAN.I.LLTLYRGRDDIF.PYK.R.AITSMGNVFP.AVS.SMSLISKPSE.INML.AIS..IY.K..LLVTD..E : 240
DMV    : ELIVTK.KELM.H.LDMSKGRIF.PKN.S.SVI..G.TIKPG...VNIYTTRNFE..FMV..ISG.MY.K..FLKPPEPD : 240

PPRV-N : KPSTSPGIGHFLRVFEIGLVRDLELGAPIFHMTNYLTV***NMSD***DYRSCLLAVGELKLTALCTPSETVTLSESGVPKREPLV : 320
PPRV-S : D...D.................G..P.V.............R.............S........R.....K... : 320
RPV    : ARRP.KVWHQDY....V.II.E.GV.T.G.......ELPRQPELET.M..L..S..A...LADSP.A.HYGR.GDDNKIR : 320
MV     : SKRSELSQLSMY....V.VI.NPG....V.......EQPV.N.LSN.MV.L.....A...HGEDSI.IPYQ.SG.GVSFQ : 320
CDV    : REFDT----REI......FIKRWLNDM.LLQT...MVLPKNSKAKV.TI.....T.AS..VEES..L.YHDSSGSQDGIL : 316
PDV    : ENFET----PEI......FINRWLGDM.L.QT...RIISNNSNTKI.TI.....A.AS...KES.IL.NLGDEESQNSVL : 316
DMV    : D.FE----FQAF.I..V.....VGSRE.VLQ...FMVIDEDEGLNF...S....R.A.V.VRGRP.VTKDI.GY.D..FK : 316

PPRV-N : VVILNLAGPTLGGELYSVLPTTDPTVEKLYLSSHRGIIKDNEANWVVPSTDVRDLQNKGECLVEACKTRPPSFC***NGTG***IG : 400
PPRV-S : ..................S.LM.............D.....................................S. : 400
RPV    : F.K.GVWASPADRDTLAT.SAI...LDG..ITT.....AAGT.I.A..V.RTD.QVKM.K.RL...RD...P...S.DWE : 400
MV     : L.K.GVWKSPTDMQSWVP.S.D..VIDR........V.A..Q.K.A..T.RTD.KLRMET.FQQ...GKIQAL.ENPEWA : 400
CDV    : ..T.GIFWA.PMDHIEE.I.VAH.SMK.IHITN...F...SI.T.M..ALASEKQEEQKG..ES..QRKTYPM..QASWE : 396
PDV    : ....G.F.A.HMDQ.EE.I.VAH.SI..IHITN...F...SV.T.M..ALALSEQGEQIN..RS...R.TYPM..Q.SWE : 396
DMV    : ..T.GII.GG.SNQKTEIY..I.SSI....IT......RNSK.R.S..AIRSD.KDKMEK.TQAL..S...PS..SSDWE : 396

PPRV-N : PWSEGRIPAYGVIRVSLDLASDPGVVITSVFGPLIPHLSGMDLYNNPFSRAAWLAVPPYEQSFLGMINTIGFPDRAEVMP : 480
PPRV-S : .................NS.............................V.....................N..... : 480
RPV    : .LEA........LTIK.G..DE.K.D.I.E.....T.D......TSFDGTKY..TT..LQN.A..TV..LVLEPSLKIS. : 480
MV     : .LKDN...S...LS.D.S.TVELKIK.A.G.....T.G......KSNHNNVY..TI..MKNLA..V...LEWIP.FK.S. : 480
CDV    : .FGGRQL.S..RLTLP..ASV.LQLN.SFTY..V.LNGD...Y.ES.LLNSG..TI..KDGTIS.L..KA.RG.QFT.L. : 476
PDV    : .FGDK.L.S..RLTL...VST.LSINVSVAQ..I.FNGD...Y.EGTLLNSG..TI..KNGTI..L..QASKG.QFI.T. : 476
DMV    : .LTSN.....AY.ALEIKED.GLELD...NY....I.GA...I.EG.S.NQD...I..LS..V..V..KVD.TAGFDIK. : 476

PPRV-N : HILTTEIRGPRGRCHVPIELSSRIDDDIKIGSNMVVLPTKDLRYITATYDVSRSEHAIVYYIYDTGRSSSYFYPVRLNFR : 560
PPRV-S : .....................R.V..........I...M.............R............L....Y.......K : 560
RPV    : N...LP..SGG.D.YI.TY..D.A...V.LS..L.I..SR..Q.VS....I..V......H..S...L...Y..FK.PIK : 560
MV     : YLF.VP.KEAGED..A.TY.PAEV.G.V.LS..L.I..GQ..Q.VL....T..V...V...V.SPS..F.....F..PIK : 560
CDV    : .V..FAP.ESS.N.YL..QT.QIR.R.VL.E..I.....QSI..VI....I...D......V..PI.TI..TH.F..TTK : 556
PDV    : ....FAP.ESSTD..L..QTYQIQ...VLLE..L.....QSFE.VV........D......V..PA.TV..T..F..KTK : 556
DMV    : .T...AVDYES.K.Y..V...GAK.Q.L.LE..L.......FG.V.....T..........V...A......F.F.IKA. : 556

PPRV-N : GNPLSLRIECFPWYHKVWCYHDCLIYNTITNEEVHTRGLTGIEVTCNPV-------- : 609
PPRV-S : .............R.................D.....................------- : 609
RPV    : .D.V..Q......DR.L..H.F.SVVDSG.G.Q.THI.VV..KI...G--------K : 609
MV     : .V.IE.QV...T.DQ.L..R.F.VLADSESGGHITHS.MV.MG.S.TVTREDGTNRR : 617
CDV    : .R.DF......V.DDNL..HQFYRFEAD.A.STTSVEN.VR.RFS..R--------- : 604
PDV    : .R.DI......V.DGHL..HQFYRFQLDA..STSVVEN.IR.RFS.DRLD------P : 607
DMV    : .E.IY........SRQL..H.Y.M.NS.VS..I.VVDN.VS.NMS.S---------R : 604
```

Fig. 2 Comparison of the predicted amino acid (aa) sequence of PPRV-S with PPRV-N and other morbilliviruses. Dot (.) represent the identity with PPRV-N; dash (-) denotes gap generated during alignment and the difference in aa sequence is represented by a single letter aa code. The membrane anchor domain is marked appropriately and the potential glycosylation sites are marked as bold and italics

the Asian and African lineage of PPR viruses [11]. A total of 16 cysteine residues identified in this protein of PPRV-S were similar to PPRV-N of which 12 are conserved across morbilliviruses, indicating the conservatory nature of this amino acid in maintaining the tertiary structure of this protein [5].

The paramyxovirus fusion protein normally synthesized as precursor F0, which is cleaved in to two subunits, F1 and F2 linked by a disulfide bond [4]. This cleavage is required for the virus to become fusogenic and thus infective [24]. The cleavage site in PPRV-S is RRTRR at position 104 and 108 is same as that of PPRV-N [11]. The three-glycosylation sites identified in the F2 subunit of PPRV-N were also conserved [11]. The fusion peptide sequence from position 109 to 133 is identical with that of PPRV-N barring one amino acid variation at position 110 (A → V). In the leucine zipper structure detected previously in the paramyxoviruses at position 459–480 of PPRV-S is conserved. At the cytoplasmic tail, the last 22 amino acids (position 525–546) were similar to that of PPRV-N. Here the last 15 amino acids are thought to interact with the M-protein during the virus budding process since an alteration in that domain leads to the abolishment of virion production [11, 26].

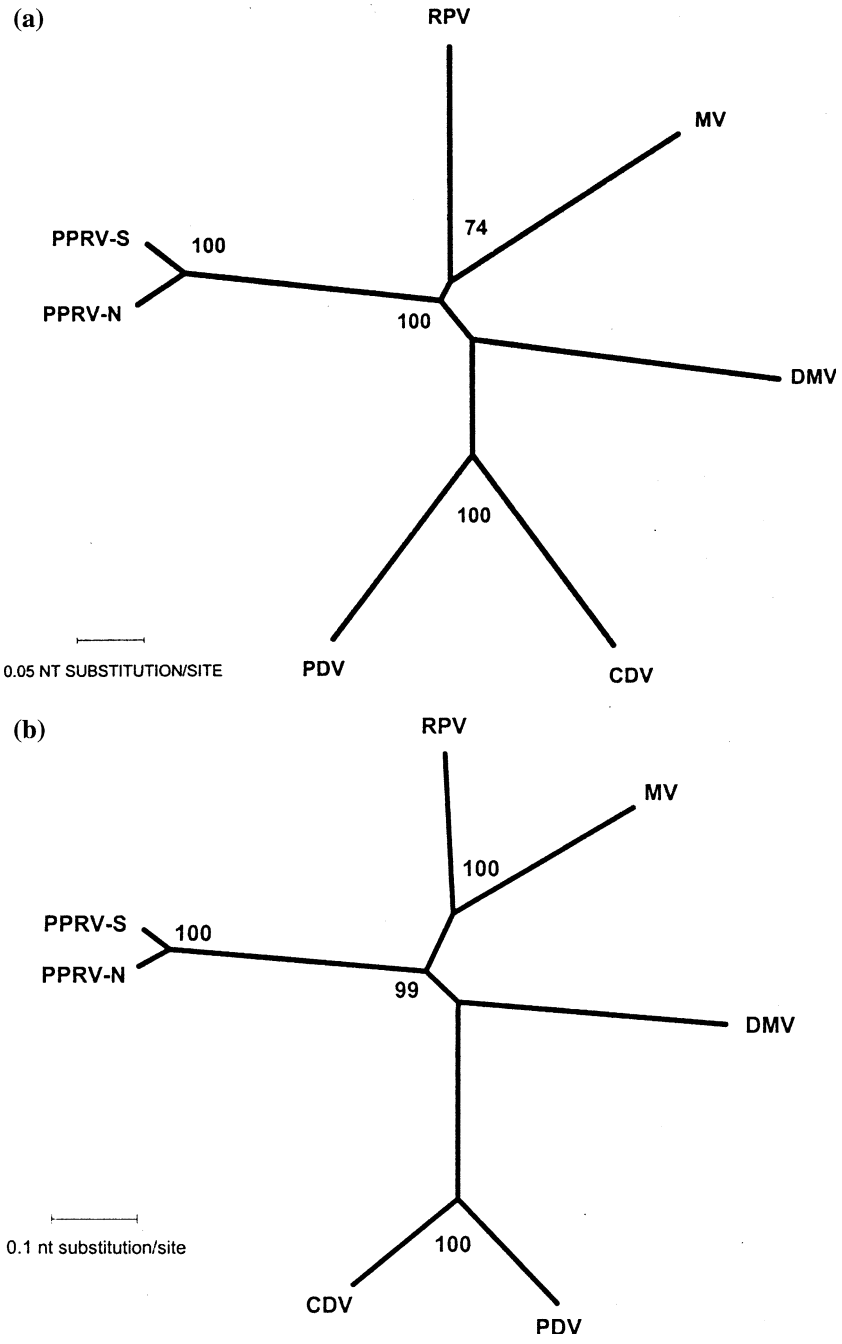

Fig. 3 Bootstrapped phylogenetic tree based on genetic distances calculated using the sequences from F gene of PPRV and other morbilliviruses. The bootstrap confidence values of major clusters are indicated in the node. The bar represents the genetic distance

Haemagglutinin protein gene

The haemagglutinin (H) protein of morbilliviruses is known to play an important role in virus attachment and induces strong neutralizing antibody response, which is highly protective [27, 28] and being the outermost protein, it is subjected to increased immunological pressure [29]. It is generally believed that the H protein of morbilliviruses lacks the neuraminidase activity, but recently it has been reported that the cytomegalovirus expressed H protein of PPRV and RPV possess neuraminidase, activity also [30].

The H gene of PPRV-S starts with usual semi conserved start signal of morbillivirus (AGGR). The complete gene along with the poly (A) is of 1954 nucleotides long, containing single ORF starting at position 18 and terminating at position 1847 with a TGA termination codon. The gene end sequence for PPRV-S is GTTAT, whereas for other morbilliviruses the sequence was ATTAT (RPV, CDV and PDV) and ATTAAG (MV). The intergenic triplet, CTT found between the H and L junction of PPRV-S is very similar to that of PPRV-N and DMV. However, MV and RPV have CGT and PDV and CDV have CTA at similar positions [31, 32].

The predicted molecular weight of H protein is ~68 kDa similar to that reported for RPV [33]. Percentage nucleotide similarity (Table 3) reveals that PPRV-S shares a homology of 90.6% with PPRV-N, but with other morbilliviruses it ranged from 33% to 45%. At amino acids level similar homology could be observed where PPRV-S had percentage amino acid identity of 92.3% with PPRV-N and with other morbilliviruses, it ranged from 34% to 49%. Similar homology also could be observed at the ORFs. This indicates the high level of divergence of the H protein among all the morbilliviruses [24, 33–35]. The H protein alignment reveals (Fig. 2) out of the 13 cysteine residues present in the protein, 12 are at identical position with all the morbilliviruses and one at position 583 is not present in the CDV and PDV. Similarly, 17 out of 37 proline residues were located at positions identical to those of other morbillivirus H proteins. The hydropathic profile (data not shown) of the predicted amino acid sequence of PPRV-S H protein was highly conserved, when compared with all other morbilliviruses [24, 33–35].

The H protein of PPRV-S is type II glycoprotein with an N-terminal proximal anchor (residues 35–58) similar to other morbilliviruses. Potential sites for asparagine (N)-linked glycosylation were found at four positions (Fig. 2) $N_{172}KSK_{175}$, $N_{215}VSS_{218}$ $N_{279}MSD_{282}$ and $N_{215}VSS_{218}$ as predicted by ScanProsite programme [36]. PPRV-N shares all of these and contains one more potential site at $N_{18}KTH_{21}$.

Phylogenetic analysis of the F and H gene

The phylogenetic analysis based on F and H gene (Fig. 3a and b) revealed that PPRV-S and PPRV-N were cluster in one group, whereas RPV and MV were grouped in a separate cluster. CDV and PDV formed another cluster and DMV was separated in a different cluster with high bootstrap confidence. The tree also confirms similar grouping of morbilliviruses as observed earlier [5, 8, 11].

Acknowledgements The authors thank Director and Joint Director (Academic), Indian Veterinary research Institute (IVRI), Izatnagar,Uttar Pradesh, India for providing necessary facilities. D.M thanks the Director, Central Institute of Fisheries Technology (CIFT), Kerala for providing financial assistance for his Doctoral Degree.

References

1. A. Haffar, G. Libeau, A. Mousso, M. Cecile, A. Diallo, Virus Res. **64**, 69–75 (1999)
2. M.S. Shaila, V. Purushothaman, D. Bhavasar, K. Venugopal, R.A. Venkatesan, Vet. Rec. **125**, 602 (1989)
3. R.P. Singh, B.P. Sreenivasa, P. Dhar, S.K. Bandyopadhyay, Arch. Virol. **149**, 155–172 (2004a)
4. R.A. Lamb, D. Kolakofsky in *Paramyxoviridae: The Viruses and Their Replication in Fields Virology*, 3rd edn., ed. by D.M. Knipe, P.M. Howley (Lippincott Williams and Wilkins, Philadelphia 2001), pp. 1305–1340
5. T. Barrett, C. Amarel-Doel, R.P. Kitching, A. Gusev, Rev. Sci. Tech. **12**, 865–872 (1993)
6. M. Mahapatra, S. Parida, B.G. Egziabher, A. Diallo, T. Barrett, Virus Res. **96**, 85–98 (2003)
7. L. Haas, M.D. Baron, B. Leiss, T. Barrett, Vet. Microbiol. **44**, 299–306 (1995)
8. A. Diallo, T. Barrett, M. Barbron, G. Meyer, P.C. Lefevre, J. Gen. Virol. **75**, 233–237 (1994)
9. M.D. Baron, T. Barrett, J. Gen. Virol. **76**, 593–602 (1995)
10. A. Diallo, Vet. Microbiol. **23**, 155–163 (1990)
11. G. Meyer, A. Diallo, Virus Res. **37**, 23–35 (1995)
12. P. Dhar, B.P. Sreenivasa, T. Barrett, M. Corteyn, R.P. Singh, S.K. Bandyopadhyay, Vet. Microbiol. **88**, 153–159 (2002)
13. M.S. Shaila, D. Shamaki, M.A. Forsyth, A. Diallo, L. Goatley, R.P. Kitching, T. Barrett, Virus Res. **43**, 149–153 (1996)
14. R.P. Singh, B.P. Sreenivasa, P. Dhar, L.C. Shah, S.K. Bandyopadhya, Vet. Microbiol. **98**, 3–15 (2004b)
15. J. Sarkar, B.P. Sreenivasa, R.P. Singh, P. Dhar, S.K. Bandyopadhyay, Vaccine **21**, 4728–4735 (2003)
16. M. Forsyth, T. Barrett, Virus Res. **39**, 151–163 (1995)
17. S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei MEGA2: Molecular Evolutionary Genetics Analysis Software (Arizona State University Tempe Arizona USA, 2001)
18. J. Felsenstein, Evolution **39**, 783–791 (1985)
19. S.B. Hedges, Mol. Biol. Evol. **9**, 366–369 (1992)
20. R. Buckland, C. Gerald, R. Barker, T.F. Wild, J. Gen. Virol. **68**, 1695–1703 (1987)
21. K. Tsukiyama, Y. Yoshikawa, K. Yamanouchi, Virology **164**, 523–530 (1988)
22. G. Bolt, M. Blixenkrone-Moller, E. Gottschalck, R.G. Wishaupt, M.J. Welsh, J.A. Earle, B.K. Rima, Virus Res. **34**, 291–304 (1994)

23. T. Barrett, D.K. Clarke, S.A. Evans, B.K. Rima, Virus Res. **8**, 373–386, (1987)
24. J. Kovamees, M. Blixenkrone-Moller, B. Sharma, C. Orvell, E. Norrby, J. Gen. Virol. **72**, 59–66 (1991)
25. M.D. Curran, Y.J. Lu, B.K. Rima, Arch. Virol. **126**, 159–169 (1992a)
26. A. Schmid, P. Spielhofer, R. Cattaneo, K. Baczko, V. ter Meulen, M.A. Billeter, Virology **188**, 910–915 (1992)
27. T. Yilma, D. Hsu, L. Jones, S. Owens, M. Grubman, C. Mebus, M. Yamanaka, B. Dale, Science **242**, 1058–1061 (1988)
28. T. Barrett, G.J. Belsham, S.M. Subbarao, S.A. Evans, Virology **170**, 11–18 (1989)
29. P.A. Rota, M.L. Hemphill, T. Whistler, H.L. Regnery, A.P. Kendal, J. Gen. Virol. **73**, 2737–2742 (1992)
30. S. Seth, M.S. Shaila, Virus Res. **75**, 169–177 (2001)
31. R. Cattaneo, G. Rebmann, A. Schmid, K. Baczko, V. ter Meulen, M.A. Billeter, EMBO J. **6**, 681–688 (1987)
32. M.D. Curran, D.K. Clarke, B.K. Rima, J. Gen. Virol. **72**, 443–447 (1991)
33. K. Tsukiyama, M. Sugiyama, Y. Yoshikawa, K. Yamanouchi, Virology **160**, 48–54 (1987)
34. M. Blixenkrone-Moller, G. Bolt, T.D. Jensen, T. Harder, V. Svansson, Virus Res. **40**, 47–55 (1996)
35. M.D. Curran, D. O'Loan, S. Kennedy, B.K. Rima, J. Gen. Virol. **73**, 1189–1194 (1992b)
36. A. Gattiker, E. Gasteiger, A. Bairoch, Appl. Bioinform. **1**, 107–108 (2002)