Infection, Genetics and Evolution 27 (2014) 163–172

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Rabies virus isolates of India – Simultaneous existence of two distinct evolutionary lineages

R.V. Chandrasekhar Reddy [a,b], B. Mohana Subramanian [c], K.S.N.L. Surendra [a], R.P. Aravindh Babu [c,d], S.K. Rana [a], K. Sunitha Manjari [e], V.A. Srinivasan [f,*]

[a] Research and Development Laboratory, National Dairy Development Board, c/o Indian Immunologicals Limited, Gachibowli, Hyderabad 500032, India
[b] Department of Biotechnology, Acharya Nagarjuna University, Guntur 522510, India
[c] Translational Research Platform for Veterinary Biologicals, Chennai 600051, India
[d] National Institute of Animal Biotechnology, Miyapur, Hyderabad 500049, India
[e] Bioinformatics Group, Centre for Development of Advanced Computing, Pune 411007, India
[f] Advisor (Animal Health), National Dairy Development Board, 33, Telecom nagar, Gachibowli, Hyderabad 500032, India

ARTICLE INFO

Article history:
Received 28 April 2014
Received in revised form 9 July 2014
Accepted 15 July 2014
Available online 29 July 2014

Keywords:
Rabies virus
Glyco-protein
Phylogeny
India

ABSTRACT

Rabies is a fatal viral disease of serious public health implication. The disease is enzootic in India. In the present study, thirty six rabies virus isolates were obtained from terrestrial mammals of India during 2002–2012. Ecto-domain coding region of the glycoprotein gene from all the isolates were sequenced and the phylogenetic analysis was performed in relation to the global rabies and rabies related virus isolates. The Indian isolates grouped into two distinctly separate lineages with majority of the Indian isolates in Arctic like 1 lineage and the remaining isolates in sub-continental lineage. Isolates of the two distinct lineages were identified simultaneously from the same geographical region. Time scaled phylogenetic tree indicated that the sub-continental lineage of the virus is one of the earliest clade of rabies virus that diverged from bat rabies virus. On the contrary, the Arctic-like 1 lineage of India appeared to be a more recent divergence event. The amino acid sequence comparison revealed that all the major antigenic sites were almost conserved among the Indian isolates whereas few amino acid variations could be identified around site IIa, minor site I and IV. The d_N/d_S study based on G ecto-domain is in support of the earlier reports of strong purifying selection. In conclusion, it is evident that the Indian rabies virus isolates are of two major distinct lineages with distant phylogenetic and evolutionary relationship.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Rabies, a viral disease, is widespread in all parts of India except in the island provinces, Andaman, Nicobar, and Lakshadweep (Sudarshan et al., 2007). The genome of rabies virus (RV), consist of a single stranded, non segmented, negative sense RNA. It belongs to the genotype-1 of the genus Lyssavirus and the family Rhabdoviridae (Tordo et al., 1986). RV can infect a wide range of wild and domestic mammals, and also humans. Though, the infection can be prevented by vaccination, there is no effective treatment after the manifestation of the disease. In the developing countries dogs are major vectors in spreading the virus and have been responsible for an estimated 55,000 human deaths annually (WHO, 2005). Non-immunized dogs are the main source of infection to other mammals (Tang et al., 2005).

The viral genome encodes five proteins in the order of **3′-N-P-M-G-L-5′**: Nucleoprotein (N), Phosphoprotein (P), matrix protein (M), Glycoprotein (G) and RNA dependent RNA polymerase (L) (Wunner et al., 1988). The rabies virus glycoprotein (RVG) is a type-I transmembrane protein with a trimeric structure, and is anchored over the viral envelope. It is composed of an endodomain (ENDO), a transmembrane region (TM) and an ectodomain (ECTO). The glycoprotein is involved in determining the viral tropism and pathogenicity (Lafon, 1994; Wiktor et al., 1973). It is the immuno-dominant antigen carrying both B- and T-cell antigenic sites in the ectodomain. The protein also plays important roles in receptor recognition and membrane fusion (Benmansour et al., 1991; Delagneau et al., 1981; Coulon et al., 1998; Dietzschold et al., 1983; Prehaud et al., 1988; Tuffereau et al., 1998). The rabies virus neutralizing antibodies are essentially directed against the RVG. The RVG gene encodes a 524 amino acid (AA) product with a 19-AA signal peptide that is cleaved inside endoplasmic reticulum to yield a mature G protein (RVG). The mature RVG contains

* Corresponding author. Tel.: +91 04 23000446.
E-mail address: vasrini@nddb.coop (V.A. Srinivasan).

http://dx.doi.org/10.1016/j.meegid.2014.07.014
1567-1348/© 2014 Elsevier B.V. All rights reserved.

an N-terminal ecto-domain, a 22-AA transmembrane-domain (TM) and a C-terminal endo-domain of 44 AA (Badrane et al., 2001).

Phylogenetic studies are important to address the various biological aspects of infections such as the relationship among the virus isolates, origin, spread and migration patterns of the virus. Phylogenic study of the rabies virus strains in India was initially based on just a few isolates (Jayakumar et al., 2004; Kuzmin et al., 2004) but later extended to a much larger number (Nadin-Davis et al., 2007; Nagarajan et al., 2006, 2009; Reddy et al., 2011). However, sequence information on RVG is available only for few Indian rabies virus isolates. RVG based phylogeny is indeed important considering its vital role in viral pathogenesis, neurovirulence, host adaptation, etc. (Badrane and Tordo, 2001; Badrane et al., 2001; Real et al., 2005).

In the present study, sequences of RVG ecto-domain coding region of 36 isolates which were collected from a wide geographical area across India were determined. The phylogenic relationship and evolutionary lineage of these isolates were studied from a global context. Our study provided evidence for the simultaneous prevalence of very early and recent evolutionary lineage of rabies virus strains in India.

2. Materials and methods

2.1. Samples and sample processing

Postmortem samples of brain and/or salivary gland tissues, from animals/humans which died of rabies symptoms, were used in the study. The samples were collected during 2002–2012 from different parts of India *viz.*, Andhra Pradesh (IAP), Tamil Nadu (ITN), Kerala (IKE), Karnataka (IKA), Maharashtra (IMA) and Uttar Pradesh (IUP) (Fig. 1). These isolates were sampled from seven different species of terrestrial mammals. Eighteen isolates from dogs, four from humans, seven from cattle, two each from buffaloes, horses, and goats, and one isolate from an elephant were collected. Impression smears were prepared from the tissue samples and then a 20% homogenate of the respective tissue samples was prepared in phosphate-buffered saline containing 2% horse serum. The impression smears were acetone fixed and stored at +4 °C until further use while the tissue homogenates were stored at −80 °C.

2.2. Laboratory test to identify rabies antigen/genome

Impression smears were developed with rabies N-gene specific mAb-FITC conjugate (Light Diagnostics™ Rabies DFA Reagent; Dean et al., 1996). N-gene (Nagarajan et al., 2009) and Ψ-gene specific RT-PCRs (Nagarajan et al., 2006) were performed to determine the presence of rabies viral genomes in the tissue homogenates. The virus in tissue homogenate was also amplified by intracerebral mouse inoculation to rule out the possible false negative results which might arise due to poor sample quality.

2.3. RNA extraction and RT-PCR

The samples with positive result in any of the above mentioned tests were further processed and were subjected to rabies G gene specific RT-PCR. Total RNA was extracted either from the post mortem tissue homogenates or mouse brain tissue homogenates using TRIzol® reagent (Invitrogen, USA), following the manufacturer's instructions. G gene specific primers were designed with Primer3-Plus software (http://primer3plus.com), based on the available G gene sequences of the Indian isolates (DQ255915 to DQ255943) and degenerate bases were included wherever necessary (Table 1). The coding sequence of RVG ecto-domain of all the isolates was amplified by RT-PCRs as two overlapping fragments.

One step RT-PCR (Qiagen, Germany) was performed using 500 ng of total RNA and the RT-PCR reaction was set up as per the protocol recommended by the manufacturer. The thermal profile employed for reverse transcription and PCR is as follows: One cycle of 50 °C for 30 min and polymerase activation at 94 °C for 10 min followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 45 s, and extension at 72 °C for 1 min. Finally a 10 min extension step at 72 °C was performed.

2.4. Sequencing and phylogenetic analysis

The amplicons were purified using QIAquick gel purification kit (QIAGEN, Germany) following the manufacturer's instruction. The purified products were sequenced with gene specific primers (which were used in RT-PCR amplification). Internal primers were used wherever necessary for getting complete sequence coverage from the PCR amplified fragments. Cycle sequencing was done with ABI Prism BigDye Terminator cycle sequencing ready reaction kit (v3.1; Applied Biosystems®) and the products were purified using EDTA-alcohol. Then the samples were resolved and analysed in a capillary gel using ABI XL 3130 in the Genetic analyzer (Applied Biosystems®).

Coding sequence of the RVG ecto-domain (1317 bp) was assembled from the sequence data using ClustalW v2 software for each of the isolate (Larkin et al., 2007). Apart from the 36 isolates of the present study, the sequences of 8 Indian rabies virus isolates from GenBank were used in the construction of the NJ tree. Additional 3 sequences of RV isolates from the Indian sub-continent (2 from Sri Lanka and 1 from Nepal) retrieved from the GenBank were included in this study. Representative sequences of RV isolates from South East Asian countries, South Korea, China, RV isolates of Artic–Artic like lineage, cosmopolitan lineage, RV isolates from bats, fixed strains of rabies virus (Flury-HEP, Pitman Moore) and rabies related viruses (RRV) were also included in the phylogenetic analysis. Details of the sequences were shown in the Table 2. The sequences were aligned using ClustalW and a NJ tree was plotted by MEGA version 6 (Tamura et al., 2013). The sequences were aligned using ClustalW and a NJ tree was plotted by MEGA version 6 (Tamura et al., 2013). Following multiple alignment, the Bayesian Information Criterion (BIC), maximum likelihood values and Akaike Information Criterion corrected (AICc) scores were also determined for the maximum likelihood fits based on the data specific model to generate the phylogenetic tree. The ML tree topology was evaluated using both neighbor-joining (NJ) and ML methods with 1000 and 500 bootstrap replicates respectively.

2.5. Evolutionary analysis

Molecular evolutionary rate and the divergence times were co-estimated and the Bayesian maximum clade credibility phylogenetic tree was constructed using Bayesian Markov Chain Monte Carlo (MCMC) analysis implemented in BEAST software package, v1.8.0 (Drummond et al., 2012). The GTR + I + G nucleotide substitution model was determined as the best fit based on the Akaike Information Criterion (AIC) scores (Posada, 2008). An uncorrelated lognormal relaxed clock model along with a coalescent tree prior was chosen and the input file for BEAST analysis was obtained using BEAUti software v1.8.0, where the sequences were tip dated according to the year of collection. The MCMC chains were run for a chain length of 2×10^8 and sampled at every thousand generations. The nucleotide substitution rate (substitutions/site/year) and the time to Most Recent Common Ancestor (tMRCA) values were obtained from the Tracer, v1.5. The posterior tree distributions were summarized using Tree annotator with the exclusion of the initial ten percent of trees and visualised in FigTree v1.3.1.

Fig. 1. Map of India showing the geographical locations of the different lineages of RV isolates of Indian origin (Map source: Yuvaraj et al., 2013).

Table 1
Sequences of the primers used in the G gene amplification by RT-PCR.

S. No.	Primer ID	Sequence (5′–3′)	Nucleotide position in full length rabies viral RNA[a]
1	GP1 For	CGCTGCATTTTRTCARAGT	3221–3239
2	GP1 Rev	GGAGGGCACCATTTGGTMTC	4116–4135
3	GP2 For	GATGARAGAGGCCTRTATAAG	4005–4025
4	GP2 Rev	RCCRCTYTTATATGACTCCCA	4854–4874

[a] GenBank ID for the reference sequence – M13215.

2.6. Deduced amino acid sequence comparison

The deduced amino acid sequences were analysed to identify the variations in the major and minor immuno-dominant sites among the Indian isolates and other Asian isolates. Representative amino acid sequences containing at least a single amino acid difference were used for the analysis. Similarly the potential N-glycosylation sites were also examined (Wunner et al., 1985; Prehaud et al., 1988; Benmansour et al., 1991). WINA software was used to plot a d_N and d_S analysis graph, to estimate the proportions of the

Table 2
Information on rabies and rabies related virus isolates used in the present study for Phylogenetic analysis.

S. No.	Virus reference	Genotype	Host	Place of origin	Year	Accession No.
1	IAP-R91	GT1	Buffalo	Andhra Pradesh	2002	KF150716
2	IKE-R73	GT1	Buffalo	Kerala	2004	KF150710
3	IKE-R77	GT1	Cow	Kerala	2004	KF150711
4	IKE-R78	GT1	Cow	Kerala	2004	KF150712
5	IKE-R86	GT1	Dog	Kerala	2004	KF150713
6	IKE-R87	GT1	Calf	Kerala	2004	KF150714
7	IMA-R88	GT1	Dog	Maharashtra	2004	KF150715
8	IKE-R94	GT1	Dog	Kerala	2004	KF150717
9	IKE-R97	GT1	Dog	Kerala	2004	KF150718
10	IKE-R101	GT1	Cattle	Kerala	2004	KF150719
11	IKE-R106	GT1	Dog	Kerala	2004	KF150720
12	IKE-R107	GT1	Dog	Kerala	2004	KF150721
13	IKE-R109	GT1	Dog	Kerala	2004	KF150722
14	IKE-R110	GT1	Cattle	Kerala	2004	KF150723
15	IKE-R111	GT1	Goat	Kerala	2004	KF150724
16	IKE-R114	GT1	Cow	Kerala	2004	KF150725
17	IKE-R116	GT1	Dog	Kerala	2004	KF150726
18	IKE-R121	GT1	Goat	Kerala	2004	KF150727
19	IKA-R129	GT1	Dog	Karnataka	2004	KF150728
20	IKA-R132	GT1	Dog	Karnataka	2004	KF150729
21	IKA-R142	GT1	Dog	Karnataka	2004	KF150730
22	IKA-R144	GT1	Dog	Karnataka	2004	KF150731
23	IMA-R146	GT1	Dog	Maharashtra	2004	KF150732
24	ITN-R148	GT1	Dog	Tamil Nadu	2005	KF150733
25	IKE-R154	GT1	Elephant	Kerala	2005	KF150734
26	IKE-R155	GT1	Cow	Kerala	2005	KF150735
27	IMA-R189	GT1	Human	Maharashtra	2009	KF150736
28	IAP-R190	GT1	Dog	Andhra Pradesh	2009	KF150737
29	IAP-R191	GT1	Dog	Andhra Pradesh	2009	KF150738
30	IAP-R192	GT1	Dog	Andhra Pradesh	2009	KF150739
31	IAP-R193	GT1	Human	Andhra Pradesh	2011	KF150740
32	IAP-R194	GT1	Human	Andhra Pradesh	2011	KF150741
33	IAP-R195	GT1	Human	Andhra Pradesh	2011	KF150742
34	IAP-R196	GT1	Dog	Andhra Pradesh	2012	KF150743
35	IUP-R197	GT1	Horse	Uttar Pradesh	2012	KF150744
36	IUP-R198	GT1	Horse	Uttar Pradesh	2012	KF150745
37	PV	GT1	Vaccine strain			
38	CVS	GT1	Vaccine strain			
39	India_CH_1998	GT1	Dog	India	1998	AY237121
40	9902NEP_1998	GT1	Goat	Nepal	1998	EU086154
41	H-08-1320_SRL	GT1	Human	Sri Lanka	2008	AB569299
42	94257_SRL	GT1	Dog	Sri Lanka	1986	EU086156
43	NNV-RAB-H_India	GT1	Human	India	2006	EF437215
44	Germany_Trp-India	GT1	Human	India	2005	AY956319
45	UK_from-India	GT1	Human	India	2010	GU936881
46	India_UP_EF151231	GT1	Dog	India	1999	EF151231
47	04029AFG	GT1	Dog	Afghanistan	2004	EU086128
48	CHAND03_India	GT1	Dog	India	1999	AY98747
49	India_DQ074978	GT1	Dog	India	2005	DQ074978
50	India_GQ233040	GT1	Dog	India	2001	GQ233040
51	9908CBG	GT1	Dog	Cambodia	1999	EU086130
52	9911CBG	GT1	Dog	Cambodia	1998	EU086131
53	01016VNM	GT1	Dog	Vietnam	2001	EU086159
54	9910LAO	GT1	Dog	Laos	1999	EU086152
55	9913BIR	GT1	Dog	Myanmar	1999	EU086129
56	04030PHI	GT1	Dog	Philippines	2004	EU086155
57	03003INDO	GT1	Dog	Indonesia	2003	EU086151
58	KRVR0901	GT1	Raccoon dog	South Korea	2008	GU937025
59	KRVC0802	GT1	Dog	South Korea	2008	GU937029
60	KRVB0907	GT1	Cattle	South Korea	2009	GU937030
61	SKRDG9901GY	GT1	Dog	South Korea	1999	DQ076100
62	SKRDG0204HC	GT1	Dog	South Korea	2002	DQ076093
63	HUN1-HM	GT1	Human	Hungary	2001	AF325462
64	86-1393_USA	GT1	Fox	USA	1986	GU936880
65	90RABN9341_Canada	GT1	Skunk	Canada	1990	RVU11752
66	93RABN0113_Canada	GT1	Red fox	Canada	1993	RVU11737
67	NY771_Canada	GT1	Raccoon	Canada	1995	U27215
68	92RBG1741_Canada	GT1	Skunk	Canada	1992	AF344305
69	MEX1-DG	GT1	Dog	Mexico	1991	AF325477
70	9811CHI	GT1	Dog	China	1998	EU086135
71	02050CHI	GT1	Human	China	1992	EU086145
72	05006CHI	GT1	Dog	China	2004	EU086147
73	05009CHI	GT1	Dog	China	2005	EU086150
74	FRA1-FX	GT1	Fox	France	1991	AF325461

Table 2 (*continued*)

S. No.	Virus reference	Genotype	Host	Place of origin	Year	Accession No.
75	USA8-BT	GT1	Bat	USA	1981	AF325494
76	USA7-BT	GT1	Bat	USA	1979	AF298141
77	ARG1-BT	GT1	Bat	Argentina	1991	AF325493
78	YUG1-BV	GT1	Bovine	Yugoslavia	1984	AF325463
79	POL1-RD	GT1	Raccoon dog	Poland	1985	AF325464
80	POL2-HM	GT1	Human	Poland	1985	AF325465
81	IRN1-HM	GT1	Human	Iran	1988	AF325472
82	9704ARG_Bat	GT1	Bat	Argentina	1997	EU293116
83	ABL-AUS_Bat	GT7	Bat	Australia	1997	AF006497
84	EBL1-POL_Bat	GT5	Bat	Poland	1985	AF298142
85	EBL2-HOL_Bat	GT6	Bat	Holland	1986	AF298145
86	Duv2-SAF_Bat	GT4	Bat	South Africa	1981	AF298147
87	Lag-NGA_Bat	GT2	Bat	Nigeria	1956	AF298148
88	Mok-ETH_Cat	GT3	Cat	Ethiopia	1990	U17064
89	8805CAM		Unknown	Cameroon	1988	AF325481
90	Flury-HEP		Vaccine strain			GU565704
91	Pitman_Moore		Vaccine strain			AJ871962

synonymous substitutions (d_S) and non-synonymous substitutions (d_N) for the entire RVG ecto-domain sequence of Indian isolates (Nei and Gojobori, 1986; Endo et al., 1996).

3. Results

Postmortem tissue samples were processed for rabies virus identification and isolation. Virus was detected by direct fluorescent antibody (DFA) test (Dean et al., 1996), RT-PCR and mouse inoculation test (Koprowski, 1996). Positive samples were used for nucleotide sequencing. Sequence data of RVG ecto-domain coding region was generated for 36 Indian rabies virus isolates in the present study. Two of the fixed strains of rabies virus (challenge virus standard – CVS and Pasteur virus strain – PV) from our repository were also sequenced (Table 2).

3.1. NJ tree analysis for G ecto-domain coding sequence

NJ tree indicated that the Indian isolates were of two different lineages with one of the Indian lineages (Indian lineage 1) was within the Arctic-like lineage (Fig. 2). While the Indian isolates were within the Arctic-like 1 lineage the South Korean isolates of rabies virus were in the Arctic-like 2 lineage. Another lineage is a distinctly separate sub-continental lineage (Indian lineage 2) and the sub-continental lineage contained Sri Lankan and Nepal isolates of RV apart from the Indian RV isolates. The Arctic-like 1 lineage is the predominant lineage among the Indian isolates for which sequence information is available. Thirty six of the 44 Indian isolates belonged to this lineage and an Afghan isolate (04029AFG) was also found clustered among the Indian isolates in this group. The sub-continental lineage comprised 6 Indian isolates, one representative Nepal isolate (many Nepal isolates of this lineage were reported by Pant et al., 2013) and two Sri Lankan isolates. Two Indian isolates were in cosmopolitan lineage and were closer to the fixed strains of RV.

There was no distinct pattern for the host species of isolation but isolates from the same geographical region clustered together. However, isolates of the two Indian lineages (Indian lineage 1 and 2) were found to exist almost simultaneously in the same area. All the isolates were genotype 1 rabies virus and rabies related virus has not been identified so for from the terrestrial mammals of India. The Indian RV isolates from the terrestrial mammals were phylogenetically distant from the bat isolates of RV which was reported in the Americas and other parts of the world. The isolates from other Asian countries such as China and south-east Asian countries formed a separate Asian lineage which is closer to the sub-continental lineage. The Indian lineage 1 is closer to the other Arctic and Arctic-like isolates from other parts of the world.

3.2. Time-scaled evolutionary tree

The maximum clade credibility tree using the ecto-domain sequences of G gene showed that the sub-continental lineage was the first to diverge from the bat rabies viruses (Fig. 3). On the contrary, the Arctic-like 1 lineage of India seemed to be a recent divergent event. Our inferred substitution rate for the ecto-domain of the G gene (7.5×10^{-4}; 95% HPD; 4.2×10^{-4} to 1.2×10^{-3}) was higher than that of the previous estimate of 3.9×10^{-4} (95% HPD $1.2–6.5 \times 10^{-4}$) for the complete G gene (Bourhy et al., 2008). Hence, mean substitution rate for each codon position was estimated as previously suggested (Sullivan and Joyce, 2005) and a higher rate was observed in the third codon position (2.45) compared to first (0.31) and second (0.13), indicating a strong purifying selection of RVG ecto-domain. As per the present tree, the mean age of the most recent common ancestor for rabies viruses of non-flying mammals was around 288 years, in the past which is much lesser than the previous estimates (Bourhy et al., 2008). This also could be the result of the RVG ecto-domain's strong purifying selection which can obscure the age of the viral lineages and result in a much younger estimate for the ancestor (Wertheim and Kosakovsky, 2011; Suchard and Rambaut, 2009).

3.3. Comparison of translated amino acid sequences

Previously reported antigenic sites of the ecto-domain of rabies virus G protein (Benmansour et al., 1991) were analysed and the antigenic sites were found conserved across the Indian isolates. The major antigenic site-II was 100% conserved in Indian isolates (both at site IIa and IIb). The other major antigenic site (site-III) was mostly conserved among the Indian isolates except for one amino acid variation in ITN-R148 at the residue 337 (E337D). The other minor antigenic sites such as site I, site a and site b, c (former site IV), were mostly conserved except for one prominent change at site I. All the lineage 1 Indian isolates had L at 231 within the site I, whereas most of the lineage 2 isolates had P. The amino acid variations were predominant in the region between site IIa and site IIb. The variations were also found around site I/site IV and towards the C-terminal region of the G ecto-domain.

The amino acid sequence variations identified between the lineage 1 and lineage 2 isolates are shown in Table 3. The 426 (Q426L) variation was unique for all the lineage 2 isolates. This unique

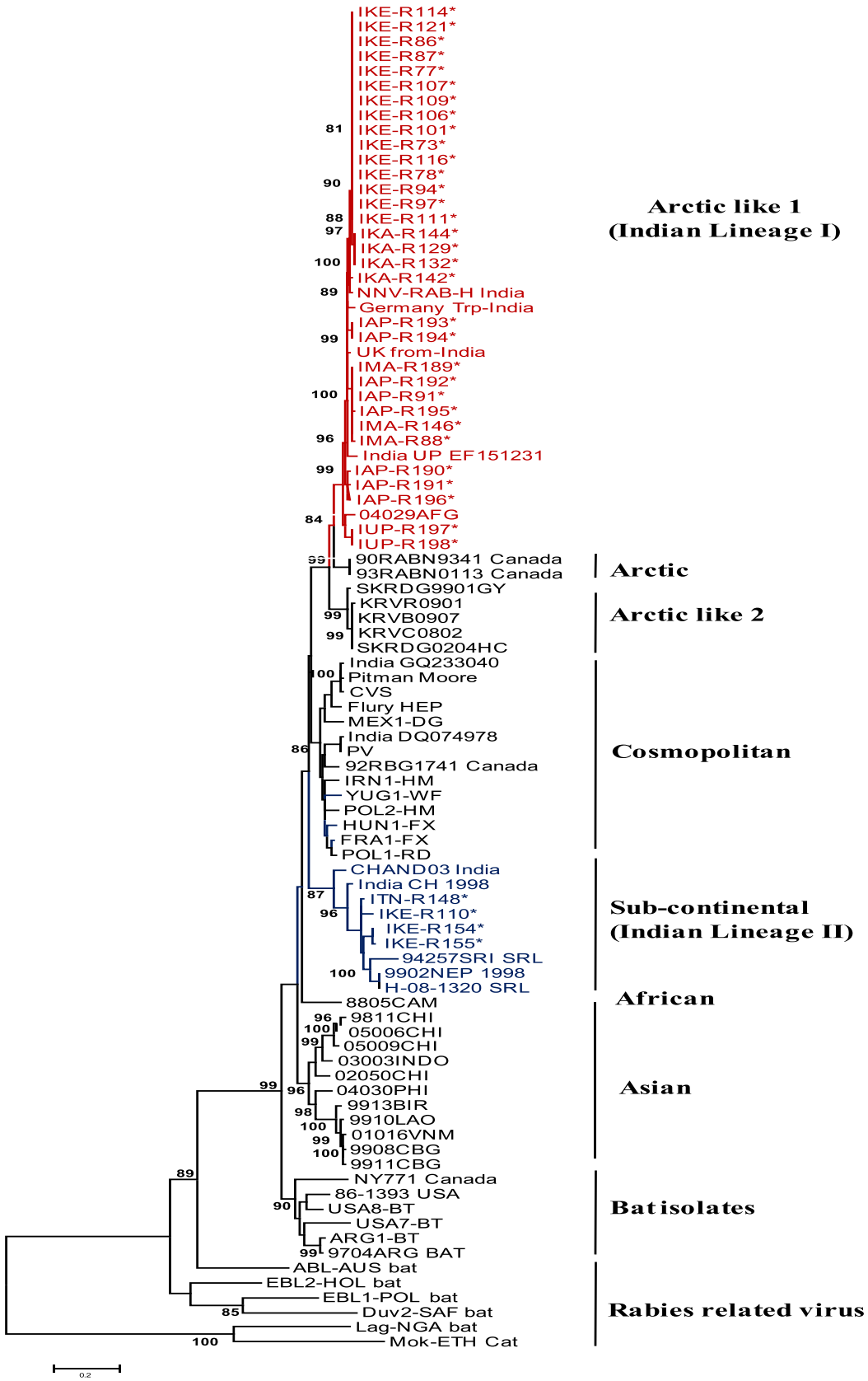

Fig. 2. Neighbor-joining tree generated by 1317 bp of glycoprotein ecto-domain showing the genetic relationship of the rabies isolates. The percentage of bootstrap values given to the left of main branch. *Isolates of the present study.

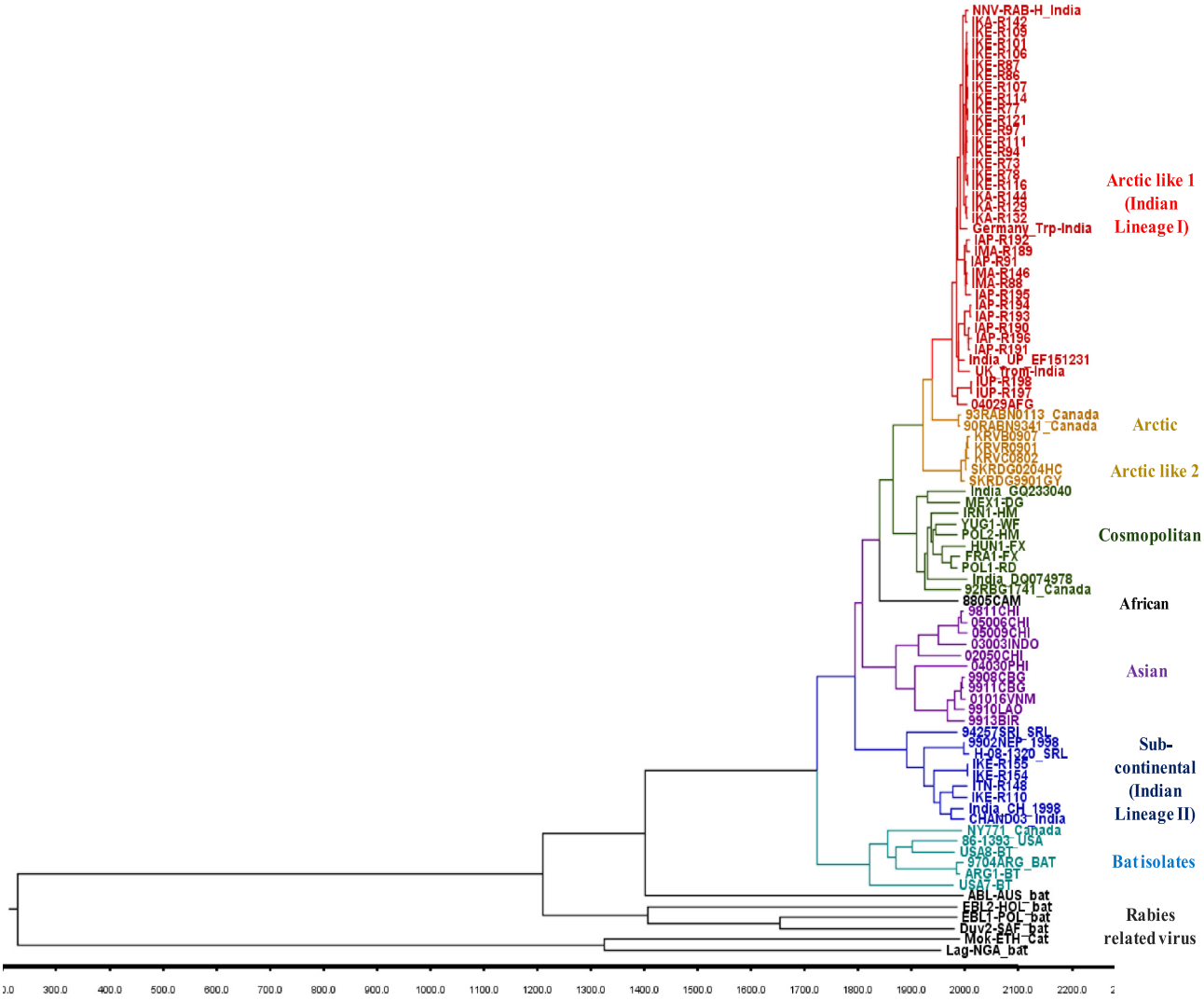

Fig. 3. BEAST tree generated based on the rabies virus glycoprotein ecto-domain gene. The time scale is shown at the bottom.

variation and the other variations in combination can be used to differentiate the two lineages of Indian isolates. The variations at 164 and 243 were common for all sub-continental isolates and other Asian isolates.

The N-glycosylation sites at amino acid positions 37 and 319 were present in all the Indian isolates. Few of the Indian isolates had one more N-glycosylation site at amino acid positions 158 or 247 and therefore, the total N-glycosylation sites of Indian isolates ranged from 2 to 3. Arginine at 333, which is considered as an essential residue for neuro-virulence of the virus was found conserved in all the isolates of the present study (Fig. 4).

3.4. d_N and d_S analysis

A d_N vs d_S analysis was performed using WINA software. The d_N/d_S ratio was plotted in a graph for the entire RVG ecto-domain sequence (Fig. 5). The d_N vs d_S analysis indicated that there was a strong purifying selection applied over the entire RVG ecto-domain region. However, many residues in the region between site IIa and IIb, and around site IV had $d_N > d_S$ and for one residue in the same region, it is $d_N > 2d_S$. Thus, the nucleotides downstream the coding region of major antigenic site IIa can tolerate non-synonymous substitutions without any deleterious effect on the virus.

Table 3
Amino acid variations between two different lineages of Indian rabies virus isolates across RVG ecto domain.

S. No.	Amino acid position in RVG	Substitutions
1	83	K → R
2	102	M → L
3	156	S → G
4	164	I → V
5	193	T → V
6	231[a]	L → P
7	243	M → I
8	426	Q → L

[a] Found within the antigenic site I of glycoprotein ecto domain.

4. Discussion

Rabies is one of the major public health threats in India and dogs are the primary source of virus spread. The studies on the viral dynamics using a substantially large number of G gene sequences of rabies virus isolates from India are sparse. RVG is the major determinant for the viral infection, antigenicity and pathogenicity and the host cell receptor recognition is mediated by RVG (Zhang et al., 2013; Benmansour et al., 1991; Thoulouze et al., 1998). Therefore, RVG based phylogeny was used for

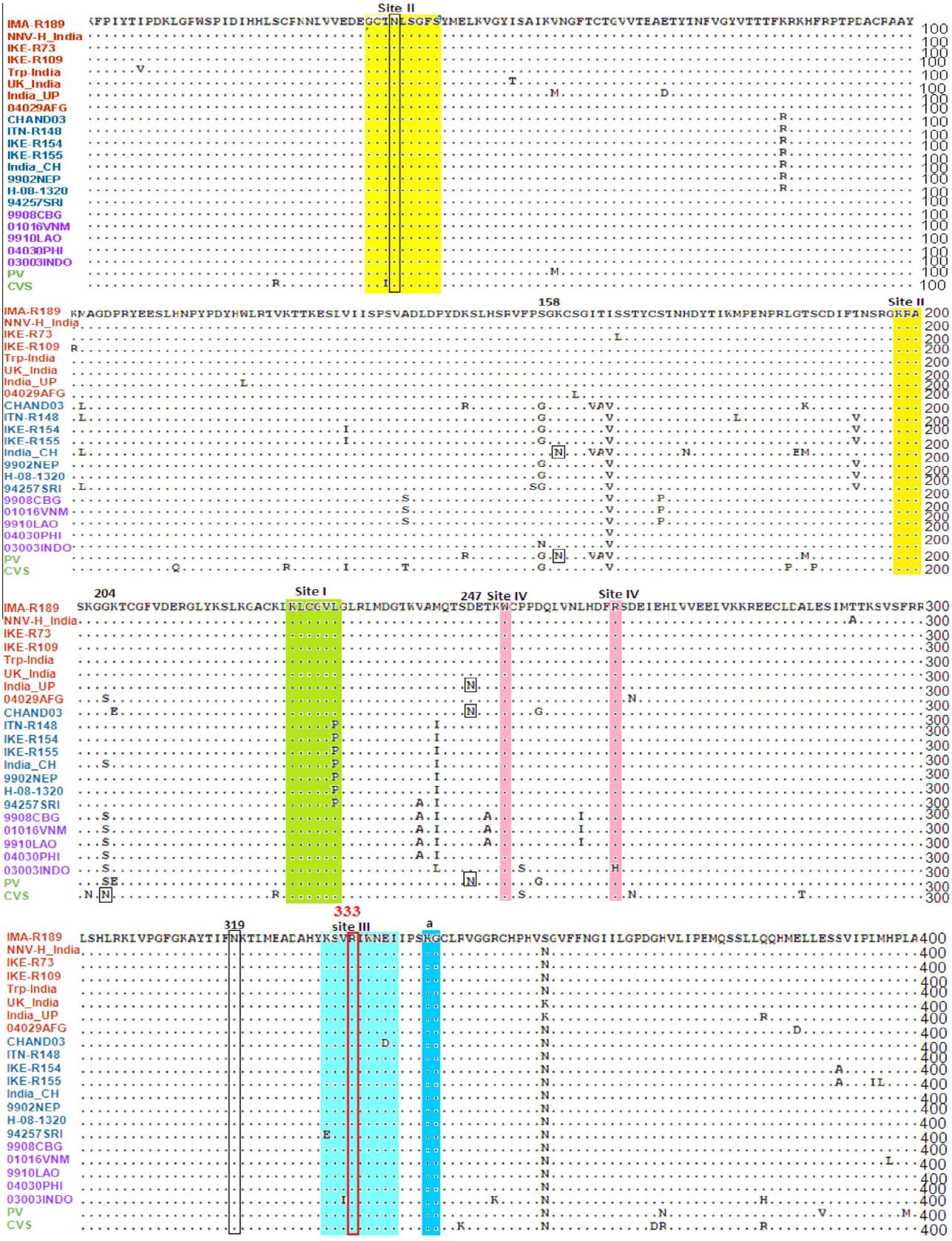

Fig. 4. Multiple alignment of partial deduced amino acid sequences of rabies virus glycoprotein. Colored boxes indicate antigenic sites; black solid boxes indicate potential N-glycosylation sites; Red box indicates Arginine 333 which is responsible for neuro-virulence. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

studying viral evolution and host adaptation (Badrane and Tordo, 2001; Meng et al., 2007). Apart from the phylogeny, translated amino acid sequence comparison and d_N vs d_S analysis was also performed in this study to rule out any antigenic variants.

In the present study, RVG ecto-domain coding region was sequenced for the 36 Indian rabies virus field isolates and their phylogenetic and evolutionary relationship were analyzed from the global context. Phylogenetically, rabies virus isolates got separated as per the established rabies virus clades such as Arctic, Arctic-like, Cosmopolitan, African, Asian, sub-continental, etc (Bourhy et al., 2008). The phylogenetic analysis identified two distinctly separate Indian lineages and the presence of two lineages among the rabies virus isolates of India was seen earlier using N gene phylogeny (Nagarajan et al., 2009; Nadin-Davis et al., 2007). The presence of virus from the Arctic like lineage (Arctic-like 1) in India was established by the earlier studies using N gene phylogeny and only two Indian isolates were identified with the sub-continental lineage (Nadin-Davis et al., 2007; Kuzmin et al., 2008) so far.

Majority of the Indian isolates of the present study were in Arctic-like 1 lineage and it is distributed all over India. Virus of this lineage is predominant in India and all the five north Indian isolates of the present study belonged to the Arctic-like 1 lineage. Five of the south Indian isolates and one north Indian isolate (obtained from GenBank) were part of the sub-continental lineage. However, sub-continental lineage was identified in a wide geographical region which spreads across Sri Lanka, India and Nepal (Pant et al., 2013). Thus, more isolates of sub-continental lineage can be identified if many north Indian samples are screened. The Arctic-like 1 lineage was reported from other sub-continental neighbors of India like Nepal, Bangladesh and Bhutan (Pant et al., 2013; Jamil et al., 2012; Reddy et al., 2011). The Arctic-like 1 lineage is the predominant lineage in the sub-continent region with the exception of Sri Lanka. All the reported isolates of Sri Lanka belonged to sub-continental lineage. A detailed analysis of the circulating virus isolates of Nepal revealed the simultaneous existence of both the lineages (Arctic-like and sub-continental) of rabies virus in the country which is similar to the Indian scenario (Pant et al., 2013). Therefore, the existence of sub-continental lineage in Bangladesh and Bhutan, and similarly the presence of Arctic-like lineage in Sri Lanka cannot be ruled out unless a large number of isolates are studied. In contrast, the Himalayan range has been acting as a barrier in preventing the spread of these lineages to China where Asian and cosmopolitan lineages were reported (Meng et al., 2007). Two of the Indian isolates grouped in cosmopolitan lineage and these isolates had very close sequence identity with the vaccine strains. In the NJ tree, these isolates grouped along with vaccine strains and therefore, these may be a result of laboratory cross contamination of vaccine viruses.

Many RV isolates belonging to the sub-continental lineage were identified in this study and these isolates were at the basal position among the rabies viruses of terrestrial mammals in the time scaled phylogenetic tree. These results indicate the probable origin of rabies virus of terrestrial mammals from this region. This observation is similar to that of Bourhy et al. (2008) who derived similar conclusion using the N gene phylogeny. However, their time scaled phylogenetic tree using a complete G gene data set indicated that the Asian lineage evolved first from the bat viruses, while the rest of the rabies lineages of the non-flying mammals evolved from the sub-continental lineage, albeit with a single isolate of sub-continental lineage in their analysis (Bourhy et al., 2008). Despite these observations, either bat rabies virus isolates or rabies virus isolates of terrestrial mammals which had close phylogenetic relationship with the bat rabies virus isolates were not reported yet from this region. A specific sample survey to identify rabies virus from bats of this region might provide better understanding on its prevalence.

As per the time scaled phylogenetic tree, it appears that the Arctic-like 1 clade of India is a more recent divergence event from the Arctic clade. The result supports the theory that the viruses of Arctic clade had descended towards south from fox to dogs instead of the northerly spread by species jump from dogs to fox (Nadin-Davis et al., 2007). This is in contrast to the N gene based time scale tree which suggested that the Arctic variants had probably evolved via a northerly spread (Kuzmin et al., 2008). However, these authors had used very few isolates (three in total) of Arctic-like 1 lineage viruses in their time scaled tree.

Thus, this region simultaneously has viruses from one of the early divergent events of the rabies virus towards terrestrial mammals and also the viruses with very recent divergent event. Additionally, rabies virus isolates of these two distinct lineages were identified in almost similar period from the same geographical

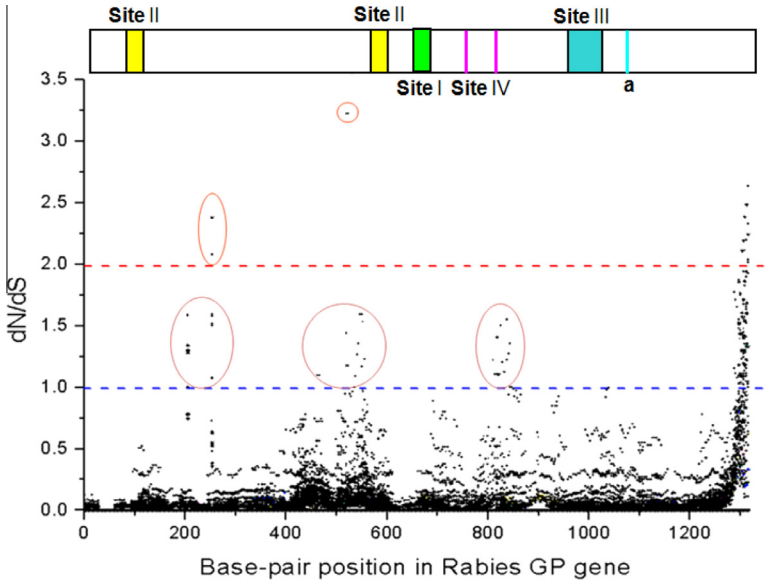

Fig. 5. Ratios of d_N to d_S along the G gene ecto-domain region. Threshold lines of the significance of the ratios are shown at values 1 and 2. A schematic representation of the G gene shows the different domains, SP, TM, ENDO, and ECTO, where antigenic sites are indicated with vertical black boxes.

region (one of the south Indian province of Kerala). Future studies might provide the status of possible recombination events between the lineages. However, the possibility of simultaneous infection with two different rabies virus isolates is very rare as the infection is 100% fatal. Additionally, the rabies viral RNA is closely associated with N protein which could prevent recombination events.

The overall d_N/d_S was very low as reported earlier for the rabies virus (Bourhy et al., 2008) and a strong purifying selection seems to be prevailing at majority of the G ecto-domain region. However, the region around the site IIa had many residues with $d_N > d_S$ and a similar observation was reported by Badrane et al. (2001) for rabies virus of carnivores. In the translated amino acid sequence comparison, all major antigenic sites were mostly conserved but variations were seen around site IIa, and minor sites I and IV. Experimental verification is necessary to find out whether these variations can affect the binding of the neutralizing antibodies with the major antigenic sites of rabies virus.

In conclusion, it is evident that the Indian rabies virus isolates are of two major distinct lineages with distant phylogenetic and evolutionary linkages. RV isolates in India are mostly canine variants and mass vaccination of dogs and intensifying public awareness about the disease are some of the ways of controlling the disease.

Acknowledgement

We thank Dr. N.M. Ponnanna, Scientist, National Dairy Development Board, C/o Indian Immunologicals Limited for critical reading of the manuscript. We also thank Dr. Ravindra Kumar Sisodiya, Dr. Shailendra Sahu and Dr. Gontu Abhinay, veterinary surgeons, Indian Immunologicals Limited, for their excellent technical support in processing suspected RV infected brain samples for this study.

References

Badrane, H., Bahloul, C., Perrin, P., Tordo, N., 2001. Evidence of two lyssavirus phylogroups with distinct pathogenicity and immunogenicity. J. Virol. 75, 3268–3276.
Badrane, H., Tordo, N., 2001. Host switching in lyssavirus history from the Chiroptera to the Carnivora orders. J. Virol. 75, 8096–8104.
Benmansour, A., Leblois, H., Coulon, P., Tuffereau, C., Gaudin, Y., Flamand, A., Lafay, F., 1991. Antigenicity of rabies virus glycoprotein. J. Virol. 65, 4198–4203.
Bourhy, H., Reynes, J.M., Dunham, E.J., Dacheux, L., Larrous, F., Huong, V.T., Xu, G., Yan, J., Miranda, M.E., Holmes, E.C., 2008. The origin and phylogeography of dog rabies virus. J. Gen. Virol. 89, 2673–2681.
Coulon, P., Ternaux, J.P., Flamand, A., Tuffereau, C., 1998. An avirulent mutant of rabies virus is unable to infect motoneurons in vivo and in vitro. J. Virol. 72, 273–278.
Dean, D.J., Abelseth, M.K., Atanasiu, P., 1996. The fluorescent antibody test. In: Meslin, F.X., Kaplan, M.M., Koprowski, H. (Eds.), Laboratory Techniques in Rabies, 4th ed. World Health Organization, Geneva, Switzerland, pp. 88–95.
Delagneau, J.F., Perrin, P., Atanasiu, P., 1981. Structure of rabies virus: spacial relationships of the proteins G, M1, M2 and N. Ann. Virol. (Inst. Pasteur) 132E, 473–493.
Dietzschold, B., Wunner, W.H., Wiktor, T.J., Lopes, A.D., Lafon, M., Smith, C.L., Koprowski, H., 1983. Characterization of an antigenic determinant of the glycoprotein that correlates with pathogenicity of rabies virus. Proc. Natl. Acad. Sci. U.S.A. 80, 70–74.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.
Endo, T., Ikeo, K., Gojobori, T., 1996. Large-scale search for genes on which positive selection may operate. Mol. Biol. Evol. 13, 685–690.
Jamil, K.M., Ahmed, K., Hossain, M., Matsumoto, T., Ali, M.A., Hossain, S., Shakhawat, H., Islam, A., Nasiruddin, M., Nishizono, A., 2012. Arctic-like rabies virus, Bangladesh. Emerg. Infect. Dis. 18, 2021–2024.
Jayakumar, R., Tirumurugaan, K.G., Ganga, G., Kumanan, K., Mahalinga Nainar, A., 2004. Characterization of nucleoprotein gene sequence of an Indian isolate of rabies virus. Acta Virol. 48, 47–50.
Koprowski, H., 1996. Mouse inoculation test. In: Meslin, F.X., Kaplan, M.M., Koprowski, H. (Eds.), Laboratory Techniques in Rabies, 4th ed. World Health Organization, Geneva, Switzerland, pp. 80–86.
Kuzmin, I.V., Botvinkin, A.D., McElhinney, L.M., Smith, J.S., Orciari, L.A., 2004. Molecular epidemiology of terrestrial rabies in the former Soviet Union. J. Wildl. Dis. 40, 617–631.
Kuzmin, I.V., Hughes, G.J., Botvinkin, A.D., Gribencha, S.G., Rupprecht, C.E., 2008. Arctic and arctic-like rabies viruses: distribution, phylogeny and evolutionary history. Epidemiol. Infect. 136, 509–519.
Lafon, M., 1994. Immunobiology of lyssaviruses: the basis for immunoprotection. In: Rupprecht, C.E., Dietzschold, B., Koprowski, H. (Eds.), Lyssaviruses. Springer-Verlag, Berlin, Germany, pp. 145–160.
Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.
Meng, S.L., Yan, J.X., Xu, G.L., Nadin-Davis, S.A., Ming, P.G., Liu, S.Y., Wu, J., Ming, H.T., Zhu, F.C., Zhou, D.J., Xiao, Q.Y., Dong, G.M., Yang, X.M., 2007. A molecular epidemiological study targeting the glycoprotein gene of rabies virus isolates from China. Virus Res. 124, 125–138.
Nadin-Davis, S.A., Turner, G., Paul, J.P., Madhusudana, S.N., Wandeler, A.I., 2007. Emergence of arctic-like lineage in India. Emerg. Infect. Dis. 13, 111–116.
Nagarajan, T., Mohanasubramanian, B., Seshagiri, E.V., Nagendrakumar, S.B., Saseendranath, M.R., Satyanarayana, M.L., Thiagarajan, D., Rangarajan, P.N., Srinivasan, V.A., 2006. Molecular epidemiology of rabies virus isolates in India. J. Clin. Microbiol. 44, 3218–3224.
Nagarajan, T., Nagendrakumar, S.B., Mohanasubramanian, B., Rajalakshmi, S., Hanumantha, N., Ramya, R., Thiagarajan, D., Srinivasan, V.A., 2009. Phylogenetic analysis of nucleoprotein gene of dog rabies virus isolates from Southern India. Infect. Genet. Evol. 9, 976–982.
Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.
Pant, G.R., Lavenir, R., Wong, F.Y.K., Certoma, A., Larrous, F., Bhatta, D.R., Bourhy, H., 2013. Recent emergence and spread of an Arctic-related phylogenetic lineage of rabies virus in Nepal. PLoS Negl. Trop. Dis. 7, e2560.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Prehaud, C., Coulon, P., Lafay, F., Thiers, C., Flamand, A., 1988. Antigenic site II of the rabies virus glycoprotein: structure and role in viral virulence. J. Virol. 62, 1–7.
Real, L.A., Henderson, J.C., Biek, R., Snaman, J., Jack, T.L., Childs, J.E., Stahl, E., Waller, L., Tinline, R., Nadin-Davis, S.A., 2005. Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. Proc. Natl. Acad. Sci. U.S.A. 102, 12107–12111.
Reddy, G.B., Singh, R., Singh, R.P., Singh, K.P., Gupta, P.K., Desai, A., Shankar, S.K., Ramakrishnan, M.A., Verma, R., 2011. Molecular characterization of Indian rabies virus isolates by partial sequencing of nucleoprotein (N) and phosphoprotein (P) genes. Virus Genes 43, 13–17.
Suchard, M.A., Rambaut, A., 2009. Many-core algorithms for statistical phylogenetics. Bioinformatics 25, 1370–1376.
Sudarshan, M.K., Madhusudana, S.N., Mahendra, B.J., Rao, N.S., Ashwath Narayana, D.H., Abdul Rahman, S., Meslin, F., Lobo, X., Ravikumar, D., Gangaboraiah, K., 2007. Assessing the burden of human rabies in India: results of a national multicenter epidemiological survey. Int. J. Infect. Dis. 11, 29–35.
Sullivan, J., Joyce, P., 2005. Model selection in phylogenetics. Annu. Rev. Ecol. Evol. Syst. 36, 445–466.
Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.
Tang, X., Luo, M., Zhang, S., Fooks, A.R., Hu, R., Tu, C., 2005. Pivotal role of dogs in rabies transmission, China. Emerg. Infect. Dis. 11, 1970–1972.
Thoulouze, M.I., Lafage, M., Schachner, M., Hartmann, U., Cremer, H., Lafon, M., 1998. The neural cell adhesion molecule is a receptor for rabies virus. J. Virol. 72, 7181–7190.
Tordo, N., Poch, O., Ermine, A., Keith, G., Rougeon, F., 1986. Walking along the rabies genome: is the large G-L intergenic region a remnant gene? Proc. Natl. Acad. Sci. U.S.A. 83, 3914–3918.
Tuffereau, C., Benejean, J., Roque Alfonso, A.M., Flamand, A., Fishman, M.C., 1998. Neuronal cell surface molecules mediate specific binding to rabies virus glycoprotein expressed by a recombinant baculovirus on the surfaces of lepidoptera cells. J. Virol. 72, 1085–1091.
Wertheim, J.O., Kosakovsky, P.S.L., 2011. Purifying selection can obscure the ancient age of viral lineages. Mol. Biol. Evol. 28, 3355–3365.
Wiktor, T.J., Gyorgy, E., Schlumberger, H.D., Sokol, F., Koprowski, H., 1973. Antigenic properties of rabies virus components. J. Immunol. 110, 269–276.
World Health Organization, 2005. WHO expert consultation on rabies. Introduction. WHO Technical Report Series 931. World Health Organization, Geneva, p. 3.
Wunner, W.H., Dietzschold, B., Macfarlan, R., Smith, C., Golub, E., Wiktor, T., 1985. Localization of immunogenic domains on the rabies virus glycoprotein. Ann. Inst. Pasteur. 136 E, 353–362.
Wunner, W.H., Larson, J.K., Dietzschold, B., Smith, C.L., 1988. The molecular biology of rabies viruses. Rev. Infect. Dis. 10, S771–S784.
Yuvaraj, S., Madhanmohan, M., Nagendrakumar, S.B., Kumar, R., Mohana Subramanian, B., Mohapatra, J.K., Sanyal, A., Pattnaik, B., Srinivasana, V.A., 2013. Genetic and antigenic characterization of Indian foot-and-mouth disease virus serotype O isolates collected during the period 2001 to 2012. Infect. Genet. Evol. 13, 109–115.
Zhang, G., Wang, H., Mahmood, F., Fu, Z.F., 2013. Rabies virus glycoprotein is an important determinant for the induction of innate immune responses and the pathogenic mechanisms. Vet. Microbiol. 162, 601–613.