Gupta *et al. BMC Res Notes (2015) 8:581*
DOI 10.1186/s13104-015-1573-2

BMC Research Notes

RESEARCH ARTICLE Open Access

Detection of 40 bp insertion-deletion (INDEL) in mitochondrial control region among sambar (*Rusa unicolor*) populations in India

Sandeep Kumar Gupta[1*], Ajit Kumar[1], Ajay Gaur[2] and Syed Ainul Hussain[1]

Abstract

Background: The mitochondrial DNA (mtDNA) control region is extensively used in the phylogeography of species. We examined sequence variations in the mtDNA control region of sambar (*Rusa unicolor*) populations from the South, Central and North India.

Results: Most of the samples collected from the south India exhibited a 40 bp insertion in the mtDNA control region. This insertion was not observed in the North and Central Indian populations.

Conclusion: This study provided a potential marker for molecular screening and identification of sambar populations in the form of a distinct 40 bp insertion. Some populations in South India did not exhibit this insertion. It indicates that there could be an ecological barrier that might be preventing the expansion of insertion-positive sambar population.

Keywords: Sambar, Genetic variation, mtDNA control region, Insertion-deletion

Background

The sambar (*Rusa unicolor*) is the largest cervid species in Southeast Asia. In India, it is widely distributed from the Himalayan foothills to the southern limits. Seven subspecies of sambar are recognised, with *R. u. unicolor* occurring in India and Sri Lanka [1]. Despite a large distribution range, very limited information is available on genetic variations for this species.

The mitochondrial DNA (mtDNA) control region has been used extensively in studying the population genetics of wild species, for example the tiger, *Panthera tigris* [2]; wild pig, *Sus scrofa* [3]; cervids [4]; sika deer, *Cervus nippon* [5]; roe deer, *Capreolus capreolus* [6]; white-tailed deer, *Odocoileus virginianus* [7]; Chinese water deer, *Hydropotes inermis inermis* [8] and black muntjac, *Muntiacus crinifrons* [9]. In the present study, we investigated genetic variations among sambar populations of selected zones of India. Sequences of the mtDNA hypervariable region *I* (HVR-*I*) were compared among the sambar populations of the South, Central and North India for evaluation of genetic variation.

*Correspondence: skg@wii.gov.in
[1] Wildlife Institute of India, Chandrabani, Dehra Dun 248 001, India
Full list of author information is available at the end of the article

Results

Overall, the mtDNA HVR-*I* region demonstrated high variability among sambar populations, and 26 distinct haplotypes were identified (Table 1). PCR amplifications that were 40 bp longer than the expected size were obtained with DNA extracted from 23 samples from South India (Fig. 1). Three samples from South India and all 38 samples from Central and North India yielded amplifications of the expected size. It indicated that the majority of the animals of the South Indian population have a unique feature in the mtDNA control region. We obtained 543–623 bp sequences of the control region using a cervid specific primer pair. After alignment of the sequences, 23 samples of South India

BioMed Central

© 2015 Gupta et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Table 1 Table showing 26 haplotypes (RUC1-26) with number of repeat in square bracket []

Haplotype	No. of Sample	GenBank Acc. No.	31	58	65	119	120	121	122	128	130	140	141	142	151	152	153	155	156	157	158	159	161	164	170	174	175	178	179	183	188	190	191	192	197	199	200	203	209
RUC1	[1]	KF133981	G	A	C	T	T	A	G	T	G	-	-	-	G	T	C	G	T	C	T	A	A	A	A	A	A	T	A	A	C	C	T	A	T	G	C	C	A
RUC2	[4]	KF133982	.	-	.	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
RUC3	[1]	KF133983	.	-	.	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RUC4	[1]	KF133984	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.
RUC5	[1]	KF133985	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
RUC20	[2]	KF648589	A	-	A	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RUC8	[1]	KF133988	A	-	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G	G	T	T	.	.	C	A	.	T	.
RUC9	[1]	KF133989	A	-	T	C	-	-	A	A	A	G	C	T	A	C	G	A	C	T	C	G	G	G	G	.	G	C	.	G	T	T	.	.	C	A	T	T	.
RUC10	[1]	KF133990	A	.	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC11	[1]	KF133991	A	.	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC12	[1]	KF133992	A	.	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC13	[1]	KF133993	A	-	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC14	[1]	KF133994	A	-	T	C	-	-	A	A	A	G	T	T	A	.	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC15	[1]	KF133995	A	-	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC21	[1]	KF648590	A	-	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	G	.	G	.	.	.	G	T	.	C	G	.	A	T	T	G
RUC16	[1]	KF133996	A	.	T	C	T	A	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	.	.	.	G	.	G	G	T	T	.	.	.	A	T	T	.
RUC17	[24]	KF133997	A	.	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G	G	T	T	.	.	C	A	T	T	.
RUC18	[4]	KF133998	A	-	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G	G	T	T	.	.	C	A	T	T	.
RUC19	[3]	KF133999	A	-	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G	G	T	T	.	.	C	A	T	T	.
RUC6	[2]	KF133986	A	-	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G	G	T	T	.	.	C	A	T	T	.
RUC7	[1]	KF133987	A	-	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	G	G	.	G	.	G	G	T	T	.	.	C	A	.	T	.
RUC22	[1]	KF648591	A	.	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	G	G	.	G	.	G	G	T	T	.	.	C	A	T	T	.
RUC23	[2]	KF648592	A	-	T	C	-	-	A	A	A	G	C	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	G	T	.	C	G	.	A	.	T	G
RUC24	[1]	KF648593	A	-	T	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.	T	T	.	.	.	.	A	T	T	.
RUC25	[1]	KF648594	A	-	T	C	-	-	A	A	A	G	T	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G	G	T	T	.	.	C	A	T	T	.
RUC26	[1]	KF648595	A	.	T	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	.	.	.	G	.	G	G	T	T	C	.	.	A	T	T	.

exhibited a unique 40 bp insertion after nucleotide (nt) position 233 (Table 1). This insertion in a majority of the South Indian samples indicated that there is a significant genetic variation in this population. Four out of the 23 insertion-positive south Indian samples were collected from a confined population from a zoological park in Goa. The 40 bp insertion was absent in three out of eight samples collected from Kalakkad-Mundunthurai Tiger Reserve (KMTR), Tamil Nadu. The 40 bp insertion was not observed in the samples of central and north Indian populations. Therefore, the 40 bp deletion in the control region after nt position 233 appears to be a potential marker for genetic screening and identification of Central and North Indian populations. The haplotype (RUC1–RUC26) sequences were submitted to GenBank (accession numbers KF133981-99 and KF648589-95, Table 1). Haplotypes RUC1–RUC5, RUC8–RUC15, RUC20 and RUC21 were molecular signatures of the insertion-positive South Indian population. The most frequent haplotype was RUC17, which was observed in 24 individuals from Central and North India. Haplotypes RUC2 and RUC18 were observed in four individuals from the South and Central India and RUC19 in three individuals from Central India (Table 1). The pairwise distance and percentage similarity matrices showed wide ranges of 0.002–0.113 and 91.60–99.83, respectively (Table 2). We observed 62 polymorphic sites (S) among 26 haplotypes. The average number of nucleotide differences (k) and nucleotide diversity (π) were 24.00615 and 0.04479, respectively. Median-joining (MJ) network tree exhibited a distinct clustering of insertion-positive (green colour) and insertion-negative (red colour) populations (Fig. 2).

Discussion and conclusions

In several studies, insertion–deletion (INDEL) markers have been used in population genetics and forensics [10–12] as it is easy to use. In the present study, we used INDEL to describe a unique 40 bp insertion in the HVR-*I* (control) region that is found mostly in the sambar population from the South India. This unique molecular feature differentiated it from the other populations of sambar from India suggesting further investigation. It is

Table 1 continued

Haplotype	No. of Sample	GenBank Acc. No.	219	220	223	230	237	260	270	277	289	292	310	329	340	350	406	408	414	430	433	436	438	450	451	467	469	472	483	496	499	510	511	530	543	576	579	580	603
RUC1	[1]	KF133981	A	T	G	T	T	T	C	T	T	C	T	T	C	T	A	T	T	G	T	T	C	C	G	A	C	C	G	T	A	A	A	C	G	G	T	T	C
RUC2	[4]	KF133982	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.	.	.	A	.	.	.	T	.	G	.	.	A	C	G	.	.	.	A	.	.	.	.
RUC3	[1]	KF133983	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	G	.	T	A	C	G	.	.	.	A	A	.	.	.
RUC4	[1]	KF133984	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	A	.	.	.	.	.	G	.	.	A	C	G	.	.	.	A	A	.	.	T
RUC5	[1]	KF133985	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	A	.	.	.	.	.	.	G	G	T	.	.	.	C	.
RUC20	[2]	KF648589	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	T	.	G	.	T	A	C	G	.	.	.	A	A	.	.	.
RUC8	[1]	KF133988	.	.	.	.	C	.	T	.	A	T	.	A	T	C	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC9	[1]	KF133989	.	.	.	.	.	.	T	.	A	T	.	A	T	C	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC10	[1]	KF133990	.	.	T	C	.	.	T	.	A	T	.	A	.	.	.	C	C	A	.	C	T	T	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC11	[1]	KF133991	.	C	T	C	C	.	T	.	A	T	.	A	T	.	.	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	T
RUC12	[1]	KF133992	.	C	T	C	.	.	.	.	A	T	.	A	T	.	.	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC13	[1]	KF133993	.	.	T	C	.	.	.	.	A	T	C	A	T	.	.	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC14	[1]	KF133994	.	.	T	C	.	.	.	.	A	T	C	A	T	.	.	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC15	[1]	KF133995	.	.	T	C	.	.	.	.	A	T	C	A	T	.	.	C	.	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC21	[1]	KF648590	.	.	T	.	.	.	T	.	A	T	.	A	T	.	.	C	C	A	C	C	.	.	.	.	T	T	.	.	.	G	.	.	.	.	.	C	.
RUC16	[1]	KF133996	.	.	.	.	-	-	-	.	A	T	.	A	T	C	.	C	.	A	.	.	.	.	A	.	T	T	A	.	.	.	.	.	.	A	C	.	.
RUC17	[24]	KF133997	.	C	.	.	-	-	-	.	A	T	.	A	T	.	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC18	[4]	KF133998	.	.	.	.	-	-	-	.	A	T	.	A	T	.	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	.	T	.	A	C	C	.
RUC19	[3]	KF133999	.	.	.	.	-	-	-	.	A	T	.	A	T	.	G	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC6	[2]	KF133986	.	.	.	.	-	-	-	.	A	T	.	A	T	.	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC7	[1]	KF133987	.	.	.	.	-	-	-	C	A	T	.	A	T	.	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC22	[1]	KF648591	.	.	.	.	-	-	-	.	A	T	.	A	T	.	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC23	[2]	KF648592	.	.	T	C	-	-	-	.	A	.	C	A	T	.	.	C	C	A	.	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC24	[1]	KF648593	.	.	.	C	-	-	-	.	A	T	C	A	T	.	.	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC25	[1]	KF648594	.	.	.	.	-	-	-	-	-	-	-	A	T	.	.	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC26	[1]	KF648595	.	.	.	.	-	-	-	-	-	-	-	A	T	C	.	C	.	A	.	.	.	.	A	.	T	T	A	.	.	.	.	.	.	A	C	.	.

Dot represent similarity with first sequence and hyphen (-) represent gap. Numerics at top represent position of variable nucleotide. Numbering of nucleotide started from the first base of control region. Alphanumeric are the NCBI GenBank accession number of haplotype sequences. Dark shade represents the 40 bp insertion

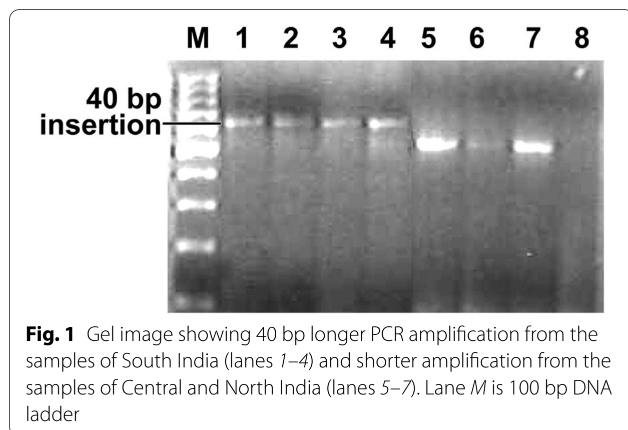

Fig. 1 Gel image showing 40 bp longer PCR amplification from the samples of South India (lanes *1–4*) and shorter amplification from the samples of Central and North India (lanes *5–7*). Lane *M* is 100 bp DNA ladder

apparent from genetic data that INDEL has a significant role in population structuring in sambar.

This study provided further insight into the genetic makeup of sambar. Analysis of samples collected from various parts of India revealed a high level of genetic variations among different populations. The percentage similarity and pairwise distance indicate that the insertion-positive haplotypes (RUC1–RUC5, RUC8–RUC15, RUC20 and RUC21) were significantly related and had a higher genetic distance from the other haplotypes (Table 2). We suspect that there could be a possible ecological barrier operating, which is separating majority of the South Indian population from the North and Central India. The 40 bp INDEL in the mtDNA region is a rapid marker for genetic

Table 2 Table showing pair wise distance (above diagonal) and percent similarity (below diagonal) amongst control region sequences generated by deer specific primer Balakrishnan et al. [10]

	RUC1	RUC2	RUC3	RUC4	RUC5	RUC20	RUC8	RUC9	RUC10	RUC11	RUC12	RUC13	RUC14	RUC15	RUC21	RUC16	RUC17	RUC18	RUC19	RUC6	RUC7	RUC22	RUC23	RUC24	RUC25	RUC26
RUC1	-	0.015	0.019	0.017	0.013	0.021	0.093	0.083	0.072	0.080	0.077	0.075	0.072	0.072	0.072	0.073	0.088	0.085	0.091	0.088	0.085	0.085	0.072	0.068	0.088	0.073
RUC2	98.38	-	0.011	0.010	0.026	0.013	0.104	0.093	0.083	0.096	0.093	0.090	0.088	0.088	0.088	0.078	0.099	0.096	0.102	0.099	0.096	0.096	0.088	0.083	0.099	0.078
RUC3	98.38	98.71	-	0.013	0.034	0.006	0.110	0.104	0.088	0.101	0.098	0.096	0.093	0.093	0.093	0.078	0.110	0.107	0.113	0.110	0.107	0.107	0.093	0.088	0.110	0.078
RUC4	98.39	98.71	98.71	-	0.028	0.015	0.102	0.096	0.090	0.093	0.096	0.093	0.090	0.090	0.090	0.080	0.099	0.104	0.102	0.110	0.099	0.099	0.090	0.085	0.102	0.080
RUC5	98.87	97.58	97.25	97.58	-	0.036	0.075	0.065	0.075	0.083	0.080	0.077	0.075	0.075	0.075	0.080	0.070	0.073	0.073	0.070	0.068	0.068	0.075	0.070	0.070	0.080
RUC20	98.22	98.55	99.52	98.55	97.09	-	0.104	0.098	0.082	0.095	0.093	0.090	0.087	0.087	0.087	0.073	0.104	0.101	0.107	0.104	0.101	0.101	0.087	0.082	0.104	0.073
RUC8	92.43	91.6	91.6	92.1	93.56	91.92	-	0.011	0.054	0.052	0.049	0.047	0.049	0.049	0.049	0.041	0.008	0.010	0.010	0.008	0.010	0.010	0.049	0.038	0.008	0.041
RUC9	93.06	92.08	91.76	92.26	94.19	92.08	99.03	-	0.049	0.047	0.045	0.043	0.045	0.045	0.045	0.032	0.004	0.006	0.006	0.004	0.006	0.002	0.045	0.034	0.004	0.032
RUC10	94.2	93.38	93.21	93.23	94.04	93.54	95.65	95.97	-	0.010	0.008	0.006	0.008	0.008	0.008	0.050	0.054	0.056	0.056	0.054	0.052	0.052	0.004	0.017	0.054	0.050
RUC11	93.56	92.41	92.25	92.9	93.4	92.57	95.65	95.97	99.04	-	0.002	0.004	0.006	0.006	0.006	0.047	0.052	0.054	0.054	0.052	0.049	0.049	0.006	0.015	0.052	0.052
RUC12	94.04	92.89	92.73	93.06	93.88	93.05	95.81	96.13	99.2	99.52	-	0.002	0.004	0.004	0.004	0.045	0.049	0.052	0.052	0.049	0.047	0.047	0.004	0.013	0.049	0.049
RUC13	93.88	93.21	92.73	92.9	93.72	93.05	95.81	96.13	99.19	99.19	99.68	-	0.002	0.002	0.002	0.047	0.047	0.049	0.049	0.047	0.045	0.045	0.002	0.011	0.047	0.047
RUC14	94.04	93.38	92.89	93.06	93.88	93.21	95.65	95.97	99.03	99.03	99.52	99.84	-	0.004	0.004	0.050	0.049	0.052	0.052	0.049	0.047	0.047	0.004	0.013	0.049	0.050
RUC15	94.04	93.38	92.89	93.06	93.88	93.21	95.65	95.97	99.03	99.03	99.52	99.84	99.68	-	0.004	0.045	0.049	0.052	0.052	0.049	0.047	0.047	0.004	0.013	0.049	0.045
RUC21	94.04	93.05	92.89	93.06	93.88	93.21	95.97	96.29	99.36	99.36	99.52	99.52	99.36	99.36	-	0.045	0.049	0.052	0.052	0.049	0.047	0.047	0.004	0.013	0.049	0.045
RUC16	94.69	93.84	94.35	94.35	94.35	94.35	96.58	96.58	96.06	95.55	96.4	96.06	95.89	96.23	96.23	-	0.036	0.034	0.038	0.036	0.038	0.034	0.045	0.038	0.036	0.008
RUC17	93.81	92.27	92.44	93.13	94.67	92.44	99.14	99.31	96.39	96.91	96.74	96.39	96.22	96.22	96.56	96.91	-	0.002	0.002	0.002	0.002	0.002	0.049	0.038	0.002	0.036
RUC18	93.98	92.6	92.77	93.29	94.49	92.77	99.14	99.31	96.21	96.73	96.56	96.39	96.21	96.21	96.56	97.07	99.83	-	0.004	0.002	0.004	0.004	0.052	0.040	0.002	0.034
RUC19	93.63	92.25	92.43	92.94	94.49	92.43	99.14	99.31	96.21	96.73	96.56	96.39	96.21	96.21	96.56	96.73	99.83	99.66	-	0.002	0.004	0.004	0.052	0.040	0.002	0.038
RUC6	93.8	92.43	92.6	93.12	94.66	92.6	99.31	99.48	96.39	96.9	96.73	96.56	96.39	96.39	96.73	96.9	99.66	99.83	99.83	-	0.002	0.002	0.049	0.038	0.002	0.036
RUC7	93.46	92.08	92.25	92.77	94.32	92.25	98.97	99.14	96.04	96.56	96.39	96.21	96.04	96.04	96.39	96.56	99.66	99.48	99.48	99.66	-	0.004	0.047	0.040	0.002	0.038
RUC22	93.81	92.1	92.44	93.13	94.85	92.44	98.97	99.48	96.56	97.08	96.91	96.56	96.39	96.39	96.74	97.08	99.83	99.66	99.66	99.83	99.48	-	0.047	0.038	0.002	0.034
RUC23	94.32	93.29	92.94	93.12	94.32	93.12	96.39	96.39	99.31	99.14	99.31	99.66	99.48	99.48	99.31	95.7	95.35	95.18	95.18	95.35	95.35	95.52	-	0.013	0.049	0.045
RUC24	94.66	93.63	93.29	93.8	94.66	93.46	97.42	97.42	99.14	99.31	99.83	99.66	99.83	99.83	99.48	96.39	96.39	96.21	96.21	96.39	96.04	96.56	98.62	-	0.038	0.043
RUC25	93.72	92.42	92.42	92.79	94.64	92.42	99.26	99.45	96.49	96.86	96.86	96.67	96.49	96.49	96.86	96.67	99.26	99.82	99.82	99.66	99.82	99.82	96.49	96.3	-	0.036
RUC26	94.3	93.2	93.93	93.75	93.93	93.93	96.14	96.14	95.96	95.96	95.96	95.59	95.4	95.77	96.14	99.08	96.14	96.14	95.77	95.96	95.77	96.32	95.04	95.22	96.49	-

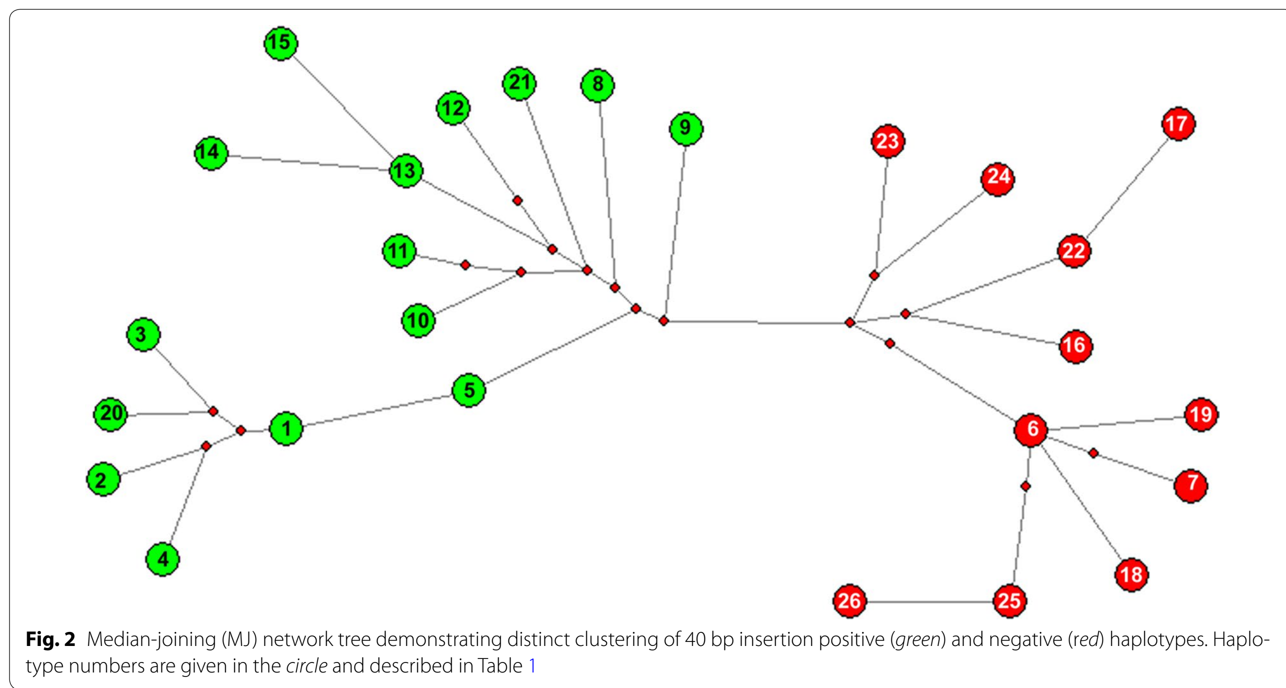

Fig. 2 Median-joining (MJ) network tree demonstrating distinct clustering of 40 bp insertion positive (*green*) and negative (*red*) haplotypes. Haplotype numbers are given in the *circle* and described in Table 1

screening and identification of these populations using a simple PCR and sequencing-based analysis.

Methods

Sampling and DNA extraction

A total of 64 biological samples (54 tissue samples and ten fecal samples) were used in this study. Of these, 26 were from South India, and 38 were from Central and North India (Fig. 3). Out of the 64 samples, 36 had been seized from poachers and forwarded by enforcement agencies to us for forensic validation. The confiscated biological samples were validated for species through sequence analysis of the mtDNA cytochrome *b* gene [13]. Biological samples were also collected from national parks and

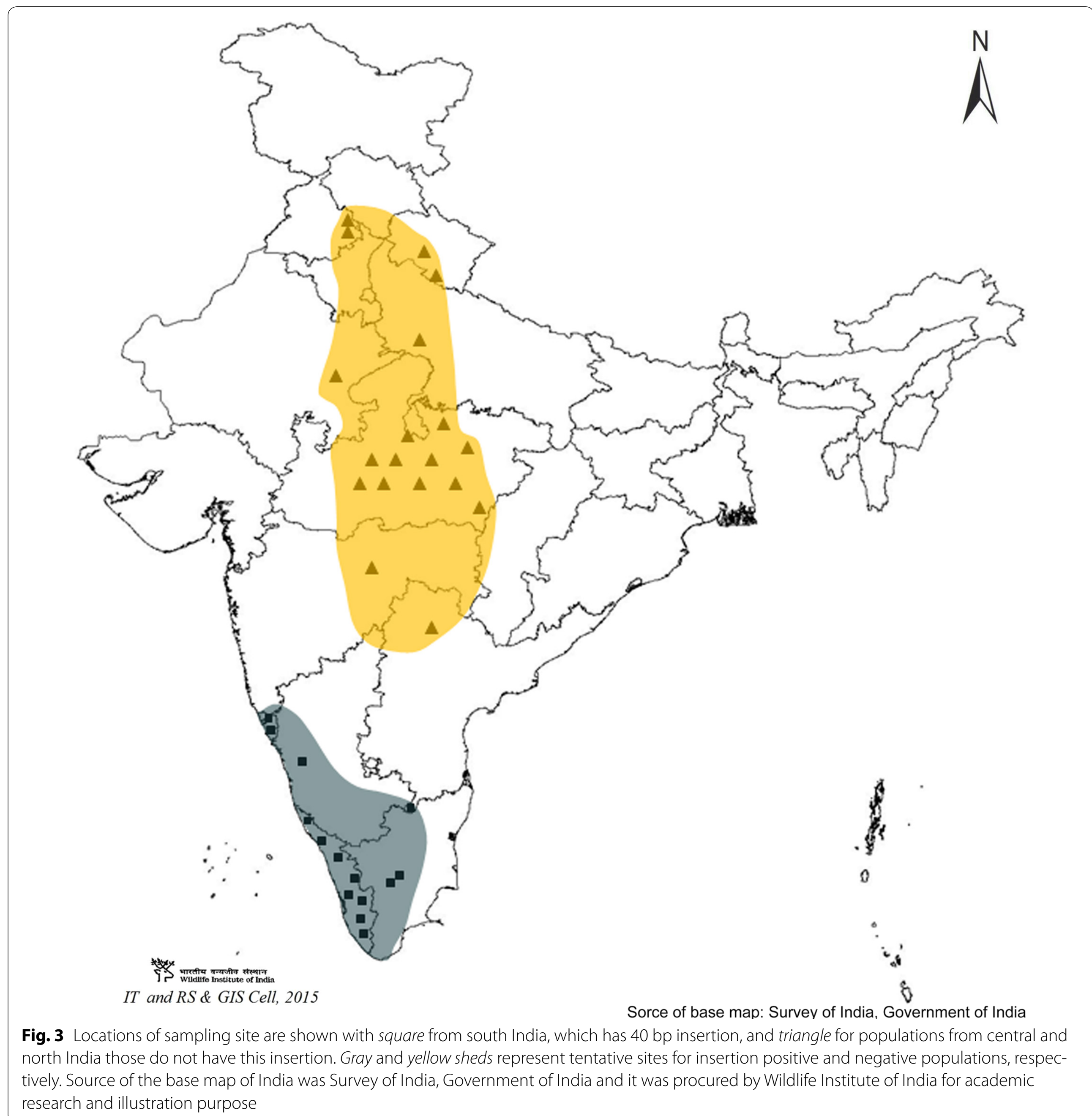

Fig. 3 Locations of sampling site are shown with *square* from south India, which has 40 bp insertion, and *triangle* for populations from central and north India those do not have this insertion. *Gray* and *yellow sheds* represent tentative sites for insertion positive and negative populations, respectively. Source of the base map of India was Survey of India, Government of India and it was procured by Wildlife Institute of India for academic research and illustration purpose

captive populations of sambar after obtaining permission from the Principal Chief Conservation of Forests (Wildlife) and Chief Wildlife Warden of the States of Uttarakhand, Madhya Pradesh, Andhra Pradesh, Kerala, Tamil Nadu, Karnataka and Goa. International regulations such as the Convention on the Trade in Endangered Species of Wild Fauna and Flora (CITES, http://www.cites.org) and Convention on Biological Diversity (CBD, http://www.cbd.int/convention) were also complied with during the study. High standards were maintained during sample collection, and no animals were harmed during this study. Six tissue samples and one shed antler sample were collected from predator kills from Panna Tiger Reserve, Madhya Pradesh. Seven and four individual's hair samples were collected from confined populations in deer parks in Andhra Pradesh and Goa, respectively. These hair samples were collected from the metal grills and concrete wall of the captive area. Ten faecal samples were

collected from forests in Kerala and Tamil Nadu. Since no animal handling was involved, approval of animal ethical committees was not required. Whole genomic DNA was extracted using phenol–chloroform [14]. A standardized protocol developed for extraction of DNA from non-invasive sampling was followed [15].

PCR amplification

The DNA extracted from above samples was used for PCR amplification of an approximately 600 bp long sequence of the mtDNA control region. A primer pair specific to the family Cervidae (deer) was used [10]. The sequences of the primer are CervtPro "CCACYATCAACACCCAAAGC" and CervCRH "GCCCTGAARAAAGAACCAGATG". PCR reactions were carried out in 20 μl reaction volumes using 1 × PCR buffer (10 mM Tris–HCl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 1 × BSA, 100 μM dNTPs, 4 pmol of each primer, 0.5 U AmpliTaq Gold DNA polymerase enzyme (Invitrogen Inc.) and 1 μl (~30 ng) of template DNA. The PCR conditions were initial denaturation at 95 °C for 10 min, followed by 35 cycles of denaturation at 95 °C for 45 sec, annealing at 54 °C for 40 sec and extension at 72 °C for 75 sec. The final extension was at 72 °C for 10 min. The efficiency and reliability of the PCR reactions were monitored using control reactions. PCR amplification was confirmed by electrophoresis on 2.2 % agarose gel stained with ethidium bromide (0.5 mg/ml) and visualised under a UV transilluminator.

DNA sequencing and analysis

The PCR products were treated with exonuclease-*I* and shrimp alkaline phosphatase for 15 min each at 37 and 80 °C. A BigDye terminator kit (version3.1) and an ABI 3130 Genetic Analyzer (Applied Biosystems) were used to generate DNA sequences from both the directions. The generated sequences were aligned by eye using Clustal W [16], available in the BioEdit package [17]. Since the primer pair used in this study amplifies approximately 40–50 bases of Proline tRNA at the beginning of the PCR reaction, the initial sequence of Proline tRNA was deleted from the aligned sequences. Hence, the start sequence examined in this study was the first nucleotide of the control region. The aligned sequences of the control region were subjected to haplotype analysis using DnaSP [18] (Table 1). A percentage similarity matrix and pairwise distance matrix were generated using Clustal W [16] and MEGA 5 [19] (Table 2). Network 4.613 software (http://www.fluxus-engineering.com) was used to generate median-joining network tree (Fig. 3).

Abbreviations

CBD: Convention on Biological Diversity; CITES: Convention on the Trade in Endangered Species of Wild Fauna and Flora; HVR-*I*: hypervariable region *I*; INDEL: insertion–deletion; KMTR: Kalakkad-Mundunthurai Tiger Reserve; MJ: median-joining; mtDNA: mitochondrial DNA; nt: nucleotide.

Authors' contributions

SKG and SAH conceived and designed this study. SKG, AK and AG carried out the sample collection. SKG and AK performed sequencing. SKG executed statistical analysis. All authors read and approved the final manuscript.

Author details

[1] Wildlife Institute of India, Chandrabani, Dehra Dun 248 001, India. [2] Laboratory for the Conservation of Endangered Species (LaCONES), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India.

Acknowledgements

This work was funded by the WII and Scientific & Engineering Research Board, Department of Science and Technology, Govt. of India. We thank the Director and Dean, WII for their support. We also thank Dr. S. Shivaji, Former Outstanding Scientist, CCMB; Dr. Y.V. Jhala, Scientist G and Prof. Qamar Qureshi, Scientist G of WII for their support. The IT and RS & GIS Cell at WII is acknowledged for their support in the preparation of map. We sincerely acknowledge the help of Forest Departments of Kerala, Madhya Pradesh, Uttarakhand, Goa, Karnataka, Andhra Pradesh, Rajasthan, Assam and Tamil Nadu States for facilitating the sample collection.

Competing interests

The authors declare that they have no competing interests.

Received: 5 March 2014 Accepted: 12 October 2015
Published online: 19 October 2015

References

1. Leslie DM. *Rusa unicolor* (Artiodactyla: Cervidae). Mamm Species. 2011;43:1–30.
2. Luo SJ, Kim JH, Johnson WE, van der Walt J, Martenson J, Yuhki N, Miquelle DG, Uphyrkina O, Goodrich JM, Quigley HB, Tilson R, Brady G, Martelli P, Subramaniam V, McDougal C, Hean S, Huang S-Q, Pan W, Karanth UK, Sunquist M, Smith JLD, O'Brien SJ. Phylogeography and genetic ancestry of tigers (*Panthera tigris*). PLoS Biol. 2004;2:2275–93.
3. Larson G, Dobney K, Albarella U, Matisoo-Smith E, Robins J, Lowden S, Rowley-Conwy P, Andersson L, Cooper A. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Sci. 2005;307:1618–21.
4. Douzery E, Randi E. The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic content. Mol Bio Evol. 1997;14:1154–66.
5. Cook CE, Wang Y, Sensabaugh G. A Mitochondrial control region and cytochromeb phylogeny of sika deer (*Cervus nippon*) and report of tandem repeats in the control region. Mol Phylo Evol. 1999;12:47–56.
6. Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G. The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. Mol Ecol. 2002;11:1285–97.
7. Moscarella RA, Aguilera M, Escalante AA. Phylogeography, population structure, and implications for conservation of white-tailed deer (*Odocoileus virginianus*) in Venezuela. J Mammal. 2003;84:1300–15.
8. Hu J, Fang SG, Wan QH. Genetic diversity of Chinese water deer (*Hydropotes inermis inermis*): implications for conservation. Biochem Genet. 2006;44:161–72.
9. Wu HL, Fang SG. Mitochondrial DNA genetic diversity of black muntjac (*Muntiacus crinifrons*), an endangered species endemic to China. Biocheml Genet. 2005;43:407–16.
10. Balakrishnan CN, Monfort SL, Gaur A, Singh L, Sorenson MD. Phylogeography and conservation genetics of Eld's deer (*Cervus eldi*). Mol Ecol. 2003;12:1–10.
11. Thangaraj K, Sridhar V, Kivisild T, Reddy AG, Chaubey G, Singh VK, Kaur S, Agarawal P, Rai A, Gupta J, Mallick CB, Kumar N, Velavan TP, Suganthan R, Udaykumar D, Kumar R, Mishra R, Khan A, Annapurna C, Singh L. Different population histories of the Mundariand Mon-Khmer-speaking Austro-Asiatic tribes inferred from the mtDNA 9-bp deletion/insertion polymorphism in Indian populations. Hum Genet. 2005;116:507–17.

12. da Costa Francez PA, Rodrigues EM, de Velasco AM, dos Santos SE. Insertion–deletion polymorphisms—utilization on forensic analysis. Int J Leg Med. 2012;126:491–6.
13. Gupta SK, Kumar A, Hussain SA. Novel primers for sequencing of the complete mitochondrial cytochrome *b* gene of ungulates using non-invasive and degraded biological samples. Conser Genet Res. 2014;6:499–501.
14. Sambrook J, Fritsch EF, Maniatis TL. Molecular cloning: a laboratory manual, vol. 2nd edn. New York: Cold Spring Harbor Press; 1989. p. 40–1.
15. Gupta SK, Kumar A, Hussain SA. Extraction of PCR-amplifiable DNA from a variety of biological samples with uniform success rate. Conser Genet Res. 2013;5:215–7.
16. Thompson JD, Higgins DG, Gibson TJ. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.
17. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 1999;41:95–8.
18. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25:1451–2.
19. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Bio Evol. 2011;28:2731–9.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit