ASSESSMENT OF GENETIC VARIATION IN SAMBAR DEER

(*Rusa unicolor*)

Thesis submitted to the

Saurashtra University, Rajkot, Gujarat

For

The Award of the Degree of

DOCTOR OF PHILOSOPHY

IN

WILDLIFE SCIENCE

By

SANDEEP KUMAR GUPTA

Ph.D. Registration No. 4513

Wildlife Institute of India

Post Box No. 18, Chandrabani

Dehra Dun - 248001, Uttarakhand, India

JULY 2014

भारतीय वन्यजीव संस्थान
Wildlife Institute of India

भारतीय वन्यजीव संस्थान
Wildlife Institute of India

Syed Ainul Hussain, *Ph.D.*
Scientist-G
Department of Landscape Level Planning & Management
Email: hussain@wii.gov.in

June 23, 2014

CERTIFICATE

This is to certify that the thesis of **Mr. Sandeep Kumar Gupta** entitled "**Assessment of genetic variation in sambar deer (*Rusa unicolor*)**" is an original piece of work submitted to the Saurashtra University, Rajkot (Gujarat), for the award of the degree of **Doctor of Philosophy** in **Wildlife Science**.

Mr. Sandeep Kumar Gupta has put in more than six terms of research work embodied in this thesis under my guidance and supervision. The work presented in this thesis has not been submitted for any other university or institution.

Hussain

[Dr S. A. Hussain]
Supervisor

Forwarded:

P. K. Mathur 23/06/2014.

[Dr. P. K. Mathur]
Dean, Faculty of Wildlife Science
Wildlife Institute of India
Chandrabani, Dehradun-248001
Uttarakhand

संकाय अध्यक्ष /Dean
भारतीय वन्यजीव संस्थान
Wildlife Institute of India
देहरादून / Dehradun

पत्रपेटी सं0 18, चन्द्रबनी, देहरादून – 248001, उत्तराखण्ड, भारत
Post Box No. 18, Chandrabani, Dehradun – 248001, Uttarakhand, INDIA
ई.पी.ए.बी.एक्स : +91-135-2640111 से 2640115 फैक्स : 0135-2640117
EPABX : +91-135-2640111 to 2640115; Fax : 0135-2640117;
ई–मेल / E-mail: wii@wii.gov.in, वेब / website: www.wii.gov.in

Saurashtra University, Rajkot

Office of the Saurashtra
University, University Road,
Rajkot – 360 005
Gujarat (INDIA)

Re-Accredited Grade "B" by NAAC
[CGPA 2.93]

Phone: +91 281 2578501

Fax: + 91 281 2586983

CERTIFICATE FOR PRE PH.D. PRESENTATION

This is to certify that **Mr. Sandeep Kumar Gupta (Regd. No. 4513, 01.01.2011)** has made Pre Ph.D. presentation as per UGC Guide Line "University Grant Commission (Minimum Standard and Procedure for award of Ph.D. Degree) Regulation–2009" and Saurashtra University Ordinance for Ph.D. Programme (O.Ph.D. 6.2), on the research work entitled **"Assessment of genetic variation in sambar deer (*Rusa unicolor*)"** in the **Wildlife Institute of India, Dehra Dun,** Research Centre of Saurashtra University, Rajkot, on **12th June 2014** before all faculty members and students of the Department for getting feedback and comments.

I also certify that the research work was appreciated by all who remain present and there was no comments made for this research work/comments made are incorporated in the thesis.

Place: **Dehradun**

Date: **12th June 2014**

Guide : **Dr. S. A. Hussain**

Department/College:

Wildlife Institute of India
Chandrabani, Dehradun – 248 001
(Uttarakhand)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF PLATES

CHAPTER 1
GENERAL INTRODUCTION

The conservation genetics is a field of amalgamation of molecular population genetics and ecology. In a broader sense, conservation genetics builds from those studies that combine population genetic data, adaptive or neutral, with data on structure of bio-geographic zones (Manel *et al.* 2003, Holderegger and Wagner 2006). Phylogeography defines the often-hostile space that separates the patches of a species habitat in a given landscape (Turner *et al.* 2001).

The recent improvements in molecular genetics tool, combined with existing or new statistical tools (geostatistics, maximum likelihood and Bayesian approaches) and powerful computers has led to the emergence of the field of conservation genetics, which is an amalgamation of molecular population genetics and ecology (Turner *et al.* 2001). This discipline aims to provide information about the intra-species interaction between bio-geographic features and micro-evolutionary processes, such as gene flow, genetic drift and selection. It will also aid in identifying cryptic boundaries, which are either breaks in the gene flow across populations without any obvious cause, or secondary contact among previously isolated populations. Population genetics can resolve population substructure of a species across different geographical scales at fine taxonomic levels (Smouse and Peakall, 1999); thus, it is different from the existing disciplines of biogeography, which focuses mainly on species diversity patterns at broad temporal and spatial scales (Brown and Lomolino, 1998), and phylogeography, which combines phylogenetics and biogeography.

Assessment of genetic variation provides a framework for directly analyzing relationships between population processes and landscape structure at relevant spatial and temporal scales. For instance, it enables researchers to test multiple competing hypotheses about the role of specific bio-geographic features and environmental conditions in affecting population connectivity. Generally, it increases our ability to make detailed inferences about movement and gene flow and potential adaptation at the bio-geographic level. These advances have proved to be exceptionally valuable in applied conservation biology. For example, understanding the bio-geographic features that drive gene flow, the spatial scales at which they act, and the temporal dynamics of

their effects on population substructure is essential to effectively use genetic data as a tool for evaluating population status and fragmentation. In addition, use of this understanding to predict, localize and implement empirically based conservation corridors would greatly improve the success of efforts to promote landscape connectivity of species at risk due to fragmentation. The potential of genetics to address large-scale connectivity questions is particularly important in the face of global climate change coupled with accelerating habitat loss and degradation.

The incorporation of the matrix into landscape (across different biogeographic zones) genetics is a discriminating difference between landscape genetics and population genetics (Holderegger and Wagner, 2008). The population genetics includes the stretches of land between occupied habitat patches as a simple function of geographical distance. In landscape genetics the atmosphere is seen as a major determinant of biological and ecological processes at the landscape level, and the different quantities and qualities of the areas that separate habitat patches are quite important (Holderegger and Wagner, 2008). For example, a patch of woodland might not hinder the movement of a ground-breeding bird found in open grasslands, but it could limit the migration of meadow butterflies or even form a complete barrier to the dispersal of meadow-plant seeds by wind (Holderegger and Wagner, 2008). Genetics across biogeographic regions does not possess its own conceptual methodological framework or its own analytical or statistical tool kit, but combines approaches and methods from bio-geographic ecology, population genetics, and spatial statistics. Hence it may be debatable that biogeographic (landscape) genetics is not a scientific discipline in itself but it provides an angle for examining spatio-temporal processes such as habitat fragmentation (Fahrig 2003). The spatial scale and extent at which biogeographic genetic research occurs are predefined by the species-specific biological and ecological process under study, and by the spatial dimension at which operational practical measures can be taken.

Biogeographic connectivity has been defined as the interaction between the movement behaviour of organisms and the structure of the various biogeographic zones (Merriam 1984, Goodwin 2003). Biogeographic connectivity thus has a structural and a functional component, and biogeographic genetics is ideally suited for testing the effect of structural bio-geographic zone connectivity (e.g., the distance between

habitat patches and the nature of the intervening habitat types) on functional landscape connectivity (i.e., dispersal and gene flow between habitat patches). Biogeographic zone ecology and population genetics address this question from different perspectives. Biogeographic zone ecology has developed a suite of tools (McGarigal 2002, Li and Wu 2004, McGarigal *et al.* 2002) for quantifying landscape patterns, and thus for measuring structural biogeographic zone connectivity.

Global extinction of species is occurring at an unprecedented rate (Pimm *et al.* 2001; Cardillo *et al.* 2006; Isaac *et al.* 2007; Morrison *et al.* 2007, Karanth *et al.* 2010). Large terrestrial mammals are the most threatened taxa with 25 per cent of species facing extinction risk and 50 per cent with declining trend (Channell and Lomolino 2000; Ceballos *et al.* 2005). South Asian mammals are among the most endangered group (Schipper *et al.* 2008). Conservation success largely depends upon identifying vulnerable species and understanding environmental factors that support their persistence, particularly in human-dominated landscapes. Geophysical factors such as climate changes and associated glaciations, as well as palaeobasins and shifting shorelines due to marine transgressions also impact on complex scenarios of species diversification (Colinvaux *et al.* 1996; Behling 2002; Antonelli and Sanmartin 2011; Aragon *et al.* 2011; Compagnucci 2011; Turchetto-Zolet *et al.* 2013). Assessment of genetic variation across various biogeographic zones would be admirable for a widely distributed species which occurs in a range of biogeographic zones.

India is the seventh largest country in the world and Asia's second largest nation. It is also a mega-diverse country with only 2.4% of the world's land area which holds 7-8% of all recorded species (Pande and Arora 2014). It is also amongst the few countries that have developed a biogeographic classification for conservation planning, and has mapped biodiversity-rich areas in the country. It is located at the tri-junction of the biodiversity rich Afro-tropical, Indo-Malayan and Palaearctic realms. Four of the 34 global biodiversity hotspots are situated in India i.e. the Himalaya, the Western Ghats, the North-east, and the Nicobar Islands. In view of the exceptional universal values and high levels of endemism in the Western Ghats, 39 sites in the States of Kerala, Karnataka, Tamil Nadu and Maharashtra have been inscribed on the United Nations Education Scientific and Cultural Organization (UNESCO) World Heritage List in 2012 (Pande and Arora 2014). The country's biodiversity is facing a

variety of threats, ranging from land use changes in natural habitats to overexploitation of natural resources, proliferation of invasive species and climate change (Pande and Arora 2014).

Species level management and assessment activities are mostly preferred among all levels of biodiversity management. Because species appeal to people, they are easily studies, and legally recognised and protected (Kumar *et al.* 2000). Genetic diversity is being increasingly considered in setting priorities for conservation of population of an endangered species. Of the 47 endemics, India ranked eighth in endemism among mammals (Kumar *et al.* 2000). The main reason for high species richness but low endemism is that the three major zoogeographical realms overlap in India. It leads to high species richness; endemism is relatively low because most species in India also occur in neighbouring countries. This is accurate for north-east India and eastern Himalaya, both has high species richness and low endemism. However, the Western Ghats are next to the above tow area in species richness has more endemics (Kumar *et al.* 2000). To examine the effect of biogeographic variations on evolutionary process a phylogeography needs to be conducted for a species having distribution in a large landscape representing different types of biogeographic conditions.

Considering above situations, sambar is one of the most suitable model ungulate for a phylogeography. It has a wide distribution range starting from Western Ghats, a more endemic area to the south-east Asian countries, a less endemic area. A phylogeography of sambar may provide insight into the origin of species and detection of ancient population among its distribution range.

1.1 SAMBAR DEER (*Rusa unicolor*): A BRIEF INTRODUCTION

Sambar (*Rusa unicolor,* Kerr 1792) is the largest oriental deer. Sambar are predominantly forest-dwellers, favoring the cover of trees, venturing out into the open mainly at night, and late at dusk or early dawn. They usually rest the whole of the daylight hours (Schaller 1967). They are good swimmers as often assemble near water (Leslie 2011). Like most deer, sambar are generally quiet, although all stags can scream a high-pitched sound when alarmed. They communicate by scent marking and foot stamping. Seven subspecies have been reported from a variety of habitats and elevations from India and Sri Lanka throughout south-eastern Asia. Body mass and

antler length have been noticed in decreasing values from west to east (Leslie 2008). Sambar populations are vulnerable because of overexploitation for subsistence and markets in meat and antlers. Sambar was upgraded in the Red-List of the International Union for Conservation of Nature and Natural Resources (IUCN) _from no status in 2006 to 'Vulnerable' in 2008 because of 50% decline in many populations over the past 3 generations (Leslie 2008). Adult sambar stags weigh between 225 and 320 kg. Sambar hinds are smaller and weigh between 135 and 225 kg (Lydekker 1916; Crandall 1964; Downes 1983.). It is closely related to the red deer (*C. elaphus elaphus*) of Asia and Europe, the russa deer (*R. timorensis*) of Asia, and the rocky mountain elk (*C. e. nelsoni*) of North America. Sambar prefers cover and avoidance of disturbance, hence the abundance of sambar may be considered as a reliable indication of the health of a forested area, and it's potential to host adequate carnivore numbers.

1.2 TAXONOMY

In Hindi, sambur and the alternate spellings sambar and sambhar are gender neutral. Latinized species names associated with sambur were masculine, or neuter. Lydekker (1915) did not present both spellings, so his use of sambar was considered a subsequent misspelling (Leslie 2011). Grubb (2005) intentionally changed the masculine gender of 11 of Heude's (1888) species names to feminine, which was warranted to concord with his combination of Heude's (1888) species names with Rusa, which he assumed to be feminine and rooted in Latin. The nomenclatural history of sambar is intricate because of its wide distribution from India through southern China and south-eastern Asia; its varied size, colour, and antler characteristics; similarity and sympatry with other Asian cervids; and the rapid pace of collection, description, and publication by Hodgson (1841) in Nepal and India and Heude (1888) in south-eastern Asia in the 1800s. Braun *et al.* (2001) tried to clarify types and type localities of many rusine deer. Earlier, Lydekker (1915) recognized 13 subspecies of *unicolor* and included *philippine nigricans, nigellus, mariannus*, and *philippinus*, and their various synonyms and name combinations, in his synonymies. Based on recent taxonomic assessments of the rusine deer (Haltenorth 1963), Grubb (2005) did not include them as synonyms of sambar. Rusine deer from the Philippines are now considered as *R. marianna* (Grubb 2005); however, Francis (2009) still considered them to be *R. unicolor*. Additional phylogenetic analyses are still needed to

clarify the phylogeny of rusine deer (Grubb and Groves 1983; Fernández and Vrba 2005).

1.3 GEOGRAPHICAL DISTRIBUTION

Rusa unicolor is the most widespread deer in Asia (Corbet and Hill 1992) and occurs from southern Nepal (Dinerstein 1979), India (Sankar 2008; Schaller 1967), Sri Lanka (Eisenburg and Lockhart 1972), and Burma (U Tun Yin 1967) throughout southern China (MacKinnon 2008; Ohtaishi and Gao 1990) and south-eastern Asia to the Pacific Coast and the islands of Borneo, Hainan, and Taiwan (Hsu and Agoramoorthy 1997; Timmins *et al.* 2008; Whitehead 1972, 1993). It occurs from sea level at various places in south-eastern Asia to about 3,000 m in the Indian Himalayas (Green 1985) and Burma (U Tun Yin 1967) and to about 3,500 m in Taiwan (Whitehead 1972).

Not many Indian ungulate has adapted itself to a wider variety of forest types and environmental conditions than the sambar (Schaller 1967). Within India, sambar occurs in the thorn forests of Gujarat and Rajasthan, in the moist deciduous forests throughout peninsular India, in the pine and oak forests at the Himalayan foothills, and in the evergreen and semi-evergreen forests of north-eastern India. Seven subspecies are reported (Timmins *et al.* 2008, Leslie 2011), occurring from Philippines in the east, to India in the west (Fig 1.1). The Indian sub-species *R. u. unicolor* is confined to India. In spite of being a common species there is little information on the phylogenetic status and population genetic structure of sambar from any parts of its distributional range.

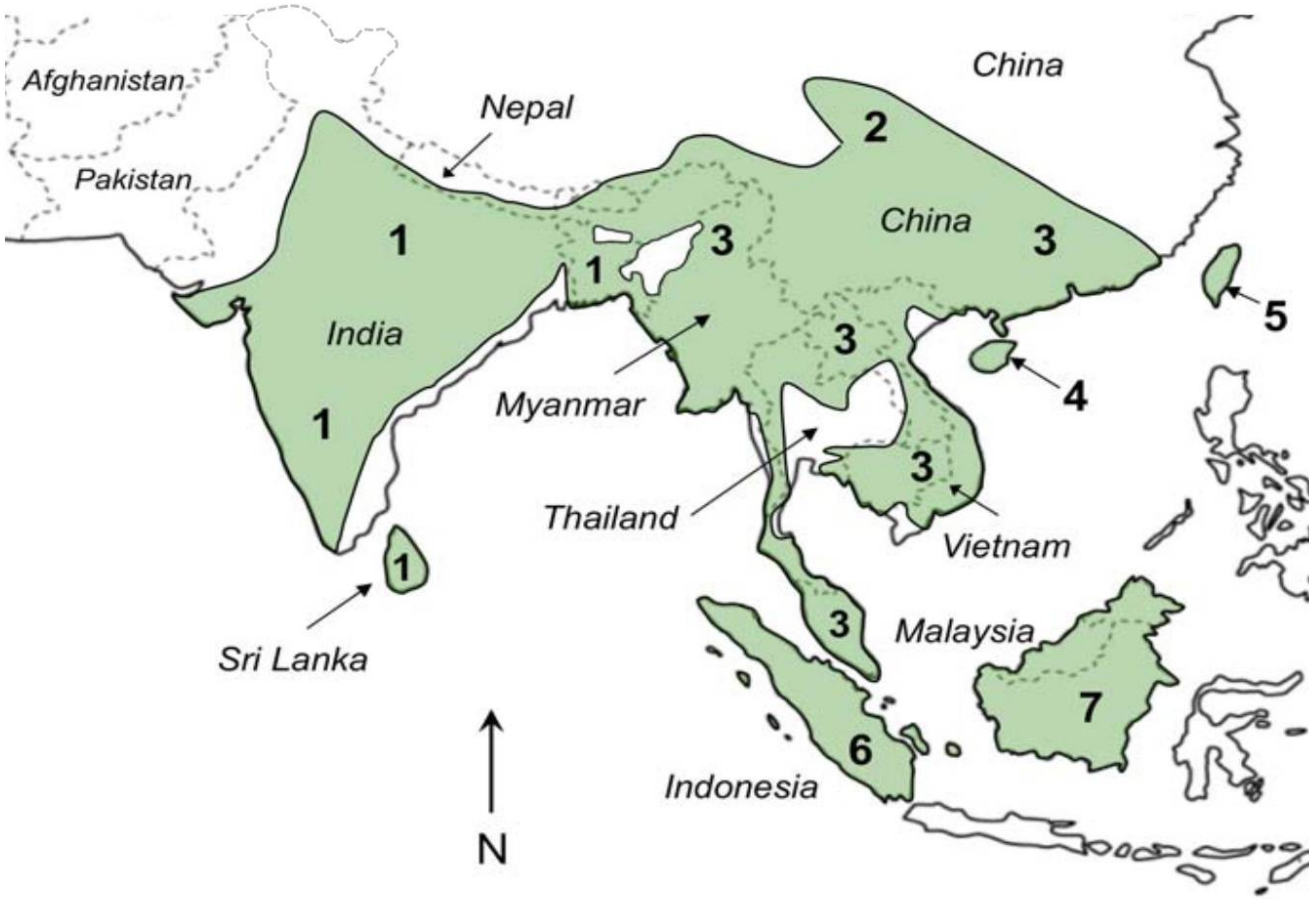


Figure 1.1: Geographical distribution of sambar in its native range (redrawn from Timmins *et al.* IUCN 2008). 1. *R. u. unicolor;* 2. *R. u. dejeani;* 3. *R. u. combojensis;* 4. *R. u. hainana;* 5. *R. u. swinhoii;* 6. *R. u. equinana;* 7. *R. u. brookei*

1.4 GENETICS OF SAMBAR

Samabr is one of the genetically least studied animals. Very few studies have been reported the genetic features of sambar (Pitra *et al.* 2004; El-Jaafari *et al*. 2008; Idris and Moin 2009). In a phylogenetic study of old world cervides, Pitra *et al.* (2004) concluded that sambar and rusa *R. timorensis* has high degree of relatedness with hog deer *Axis porcinus* with 99–100% statistical support. This association was surprising and has not been proposed before; it laid a space for further investigation. However no much study was conducted further on rusa group. El-Jaafari *et al.* (2008) studied the genetic features within and among three deer species in Malaysia, namely *Cervus nippon* (sika), *R. timorensis* and *R. unicolor* using conventional Randomly Amplified Polymorphic DNA (RAPD) markers. On the basis of polymorphisms pattern within species and between species, it was indicated that the rusa, sambar and sika deer species in Malaysia are genetically diverse. They also concluded that deer farming and establishment of population from limited founder individuals have not affected the genetic variability within the deer species in Malaysia (El-Jaafari *et al.* 2008). However, in modern molecular study use of RAPD markers is limited for animals groups and researchers discouraging the use of this marker. This is because of RAPD result are not consistent across the experimental conditions. This could be a reason that high degree of variation was observed within the farm animals. However, used of other markers like microsatellites (STR's) may yield a different outcome in this situation.

Several studies reported fertile offspring after hybridization between cervides including red deer and wapiti (Nugent *et al*. 1987), red deer and sambar deer (Muir *et al*. 1997) and sika deer crossed with red deer (Davidson 1983). In recent past, a cytogenetic study was conducted for detecting the effect of hybridization in Malaysian farmed deer industries (Idris and Moin 2009). It was based on natural mating between Bornean sambar and rusa deer and the number and chromosomal location of the nucleolar organizer regions were analyzed in their offsprings. Chromosomal banding pattern was investigated to detect chromosomal variation and to define the chromosomal homology between Bornean sambar deer (*R. u. brookei*) and rusa deer (*R. timorensis*). It was observed that rusa (chromosome no. 60) and Bornean sambar (chromosome no. 62) have different chromosome composition; however, they share the same fundamental number (70) of chromosome arm. Their hybrids were having

odd (2n = 61) chromosome number, and found with 9 metacentric to submetacentric autosomes and 24 pairs of acrocentric autosomes with two acrocentrics. One submetacentric chromosome was unpaired (Idris and Moin 2009). The morphology of the sex chromosomes in the F1 hybrids was similar to that of the parental species (Idris and Moin 2009).

1.5 ECOLOGY OF SAMBAR

1.5.1 Group size and composition

Sambar is essentially a non-social or solitary species. In sambar the typical group is small, numbering fewer than six individuals (Schaller 1967). The characteristic social unit is one hind and one fawn or one hind, one yearling and one fawn (Schaller 1967). During the rut, dominant stags are frequently seen with hinds and occasionally with other stags who may challenge the dominant stag for breeding opportunities. Average group size of Sambar is reported to be 4 to 5 individuals. In Sariska the average group size of sambar was about four individuals (Sankar 1994). Mean group size of Sambar in Pench Tiger Reserve, in central India, was 1.7 (Biswas and Sankar 2002) and in Ranthambhore Tiger Reserve, the mean group size was 3.7 (Bagchi *et al.* 2004).

1.5.2 Antler condition and breeding season

Body mass and antler length of sambar are highly variable and generally decrease from west to east across its distribution (Geist 1998; Pocock 1942). Sexes are distinguished by larger male mass (males in India 225–320 kg; females, 225 kg—Sankar and Acharya 2004), male-only antlers on short pedicles, and generally lighter color of females and young (Blanford 1888; Brander 1923; Jerdon 1874; Lydekker 1898, 1915). Across the range of sambar, head and body length is 162–246 cm, tail length is 25–30 cm, and shoulder height is 102–160 cm (Nowak 1999). Antlers of sambar are generally only 3-tined, rough, and corrugated as male's age, often robust, and consist of an anterior more or less straight brow tine that comes off at an acute angle to the main or posterior beam, which forks but once (Pocock 1942; Blanford 1888; Pocock 1933; Allen 1940; Jerdon 1874). The anterolateral tine tends to follow the main beam and is typically longer that the posteromedial tine; however exceptions have also been noticed (Groves 2003; Whitehead 1972). In few cases, the brow tine can be about 50% of total length of the main beam; 10% of males can have a 4^{th} tine

(Plate 1) on 1 antler and rarely the brow tine is bifurcate (Brander 1923; Gilbert 1888; Ward 1896; Brander 1923). The space between antlers is V-or U-shaped, and tips of tines are often inturned (Brander 1923; Downes 1983b; Lydekker 1898, 1915). Mean record antler length of sambar is 109.8 cm, exceeded by only red deer, North American elk and caribou (*Rangifer tarandus*) among cervids (Clutton-Brock *et al.* 1980; Whitehead 1993).

The period of breeding (rut) of sambar is determined by the annual antler cycle of antler development, the frequency of sexual behavior, and, in a way, the time of fawning. Sambar stags exhibited a distinct antler cycle in Sariska. Hard antlers were shed during the summer, followed by emerging and velvet antlers during monsoon months. During the remaining part of the year, sambar remained in hard antler stage. In Bandipur, sambar stags in hard antlers were observed largely between November and April and most males had shed their antlers by May (Johnsingh 1983). In India the peak rut of sambar occurs between October and December (Schaller 1967). Schaller (1967) reported that in Kanha the rut spreads over a period of at least seven months with a peak in November-December. In Sariska the peak rutting season was in winter when almost all stags were carrying hard antlers. In the wild, female sambar probably experience puberty at 18–24 months (Plotka 1999; Sheng and Ohtaishi 1993). A captive female in India reached sexual maturity at 18 months of age and gave birth at 26 months of age (Acharjyo and Misra 1971). Gestation is about 8 months, although some reports suggest that it can be longer (Hayssen *et al.* 1993, Plotka 1999; Sheng and Ohtaishi 1993).

Based on population estimate in Sri Lanka it was suggested that female may bred every alternate year (Eisenburg and Lockhart 1972). In Perak, Malaysia, only 9 of 23 females were pregnant when sampled throughout the year (Khan and Khan 1968). The number of young in a year per 100 females was typically, 50:100, even where introduced: 11–44:100 (Johnsingh 1983), 16–43:100 (Berwick 1974), 17–24:100 (Varman and Sukumar 1993), 33.7:100 (Schaller 1967) and 38.2:100 (Bagchi *et al.* 2003) in India. Female offspring may remain with their mothers as yearlings, but males leave their mothers after about 1 year (Lewis *et al.* 1990).

Photo credit: Author

Plate 1: Sambar antler collected from Jhilmil Jheel, Uttarakhand with 4th tine

1.5.3 Sex ratios

Schaller (1967) estimated a sex ratio of 0.2 males: 1 female in Kanha Tiger Reserve (TR). In Bandipur the average: female ratio in Nagarahole was 0.4:1 (Karanth and Sunquist 1992). In Sariska, the estimated male: female ratio was 0.3:1, and the female: fawn ratio was 1:0.3 (Johnsingh 1983). The male average male: female ratio was 0.1:1 and the average female: fawn ratio was 1:0.2 (Sankar 1994). In Gir, the average male: female ratio was 0.5:1, and the female: young ratio April to December and the peak fawning period was in May and June (Schaller 1967). The relatively low male numbers may be either due to selective predation, or sambar stags may be more vulnerable to stress (Leslie 2011).

1.5.4 Food habits

Dietary selection of sambar varies considerably depending on season, location, habitat variety and its effect on forage availability and quality, competitive interactions, and human activities (Kushwaha *et al.* 2004), whether in native habitat (Bagchi *et al.* 2003; Johnsingh and Sankar 1991; Schaller 1967) or introduced locations (Kelton and Skipworth 1987). In India, samabr consumes a greater variety of plants than any other ungulate (Schaller 1967), often uses cultivated areas, and is not deterred by fences as high as 2 m (Baker 1898; Brander 1923). Sambar have been observed to feed on more than 139 species of plants (Schaller 1967). The food requirements of sambar are less specialized than those of other deer (Schaller 1967). Sambar would graze or browse depending upon the forage available at any given point of time (Bentley 1978). Young green grasses are the preferred forage of sambar in Kanha TR, but browse is often important during seasons when green grasses are scarce (Schaller 1967). In Sariska TR, sambar were observed grazing as long as the green grasses are available, but switched over to browse and fallen leaves, flowers and fruits in winter and summer (Sankar 1994). Mineral licks are used regularly by sambar, particularly at night (Schaller 1967, Peacock 1933).

1.5.5 Predation and mortality

Sambar constitute one of the largest, and in turn, the most favored prey species of large carnivores such as the tiger, leopard and dhole as reported from Kanha TR, Bandipur National Park (NP) Rajaji NP and Sariska TR (Schaller 1967; Johnsingh 1983; Sankar 1994). Other than the chital (*Axis axis*), sambar are numerically the

second most important prey species of the large carnivores of India. Predation by tiger, leopard and dholes is the main cause of mortality in sambar, though they are also a favorite with hunters and poachers. In Kanha TR, sambar remains were found in nearly 11% of tiger scats and 9% of leopard scats analyzed (Schaller 1967). In Bandipur NP, their remains were found in about 30% tiger scats, 14% leopard scats, and 14% dhole scats (Johnsingh 1983). In adjoining Nagarahole NP, remains of sambar were observed in about 25% tiger scats, 13% leopard scats, and 10% dhole scats (Karanth and Sunquist 1995). In Sariska TR, around 51% of the scats of tiger and around 20% of leopard scats contained sambar remains (Sankar 1994). Sambar remains were found in nearly 14% of tiger scats in Pench TR (Biswas and Sankar 2002) and 50% of tiger scats in Ranthambhore TR (Bagchi *et al.* 2003). Mortality of sambar stags is usually high relative to their representation in the population (Johnsingh 1983, Karanth and Sunquist 1995). Males are apparently more susceptible to predation (Leslie 2011). Weakened condition after rut and territorial contests may be responsible to make males vulnerable to predation (Schaller 1967).

1.5.6 Sore spot

Sambar exhibits a unique and interesting morphological feature. Reproductively active male and females during pregnancy and lactation may have a sore spot of 10–15 cm on their throats, ventrally halfway down the neck (Geist 1998; Schaller 1967; Evans 1912). It could be a glandular part, which often exuding whitish oily or watery discharge from a blood-red spot (Davar 1938; Evans 1912; Morris 1938; Phythian-Adams 1951; Thom 1937). Thom (1937) explained the early speculation on the probable cause or function of the sore spot, which included disease (leprosy in Thailand), consumption of wild olives, ticks or some other parasite, irritation from rubbing or moving through thick coarse grass, or a wound from an attack by a species like marten, *Martes flavigula* (Davar 1938; Evans 1912; Peacock 1933; Whitehead 1972).

The sore spot was apparently not observed on sambar in Sri Lanka, among introduced populations on St. Vincent Island, Florida, and Australia, or under some captive conditions (Kurt 1978; Shea *et al.* 1990; Downes 1983; Evans 1912; Mary and Balakrishnan 1984; Thom 1937; U Tun Yin 1967). Richardson (1972) noted that all sexes and ages of introduced sambar in Texas exhibited a whirl of hair on the throat

that was relatively bare, but no secretions or blood was noticed in this population. It was summarized that perhaps this gland does not manifest itself when densities are relatively high or group size is large, as in Sri Lanka or under confinement, suggesting a role in communication (scent dispersal), particularly during breeding and at low densities (Mary and Balakrishnan 1984; Schaller 1967; Geist 1998). The sore spot was observed during breeding season in some places in India, but it was not observed during breeding, in western or northern India (Johnsingh 1980; Sankar *et al.* 2004).

1.5.7 Threat to sambar

In India, poaching has seriously depleted the abundance of large mammals in majority of the areas including high-profile areas such a Corbett Tiger Reserve and Rajaji National Park (Johnsingh *et al* 2004). There are several ways of hunting of deer in India using snares, dogs and guns and these methods are general across their range (Jathanna *et al* 2003; Kumara and Singh 2004). Two areas of moderately hunted and heavily hunted site were compared for impact of hunting in Nagarahole National Park where no significant difference in sambar densities was recorded (Madhusudan and Karanth 2002). Mostly these hunting are for village consumption of meat, but majority of its proportions in South-east Asia is sold commercially (Duckworth *et al* 1999; Kumara and Singh 2004). The commercial hunting causes major declines because the market is limitless (Steinmetz *et al* 2006). Adult males are preferred because antlers are widely displayed as trophies and are used in traditional medicine (Martin 1977). Due to excessive legal or illegal exploitation, population of sambar suffered rapid decline in neighboring countries (Tungittiplakorn and Dearden 2002; Kawanishi *et al* 2014). A recent study in Peninsular Malaysia indicated that the sambar populations are declining rapidly and they have lost more than 50% of their historical range in the past century (Kawanishi *et al.* 2014). It has been relentlessly poached for local meat consumption and now persists in only a few areas those have implemented strong enforcement. Due to the mass decline of its historical habitat, current rarity, and the lack of capacity and resources for large-scale restocking and effective protection it was suspected that the process of extinction will be exacerbated for this species in Peninsular Malaysia. Thus if these observations in Malaysia reflect global trends across the full species range, it was recommended that the IUCN Red List authority review the Red List category of sambar and may upgrade it from Vulnerable (VU) to Endangered (EN) A2cd or A4cd (Kawanishi *et al* 2014). Due to excessive harvest and

habitat loss (Timmins *et al.* 2008), sambar is now rare in Bangladesh (Basbar *et al.* 2001), Thailand (Ngampongsai 1987), Laos (Timmins and Evans 1996), and Vietnam (Khun and Kan 1991).

1.6 REVIEW OF LITERATURE

1.6.1 Phylogeography

Phylogeography is a new and rapidly developing field that characterizes the phylogenetic development of genealogical lineages across the geographic landscape (Avise *et al* 1994). It is the study of the historical processes that may be associated with the contemporary geographic distributions of individuals. It is achieved by considering the geographic distribution of individuals in view of the patterns associated with a gene genealogy (Avise 2000). The term phylogeography was used in 1987 in intraspecific phylogeography (Avise *et al* 1987). This term was coined to explain geographically structured genetic signals within and among species; and an unambiguous focus on a species biogeography or biogeographical history for phylogeography apart from conventional population genetics and phylogenetics (Knowles and Maddison 2002). Phylogeography is an approach to historical biogeography on an ecological scale of time (Ronquist 1997). It is a modern branch of biogeography in which the primary units of analysis are monophyletic clades that are inferred from phylogenetic analysis. This typically involves the use of one or more molecular markers when phylogeny is being examined, but in principle any set of phylogenetically edifying characters could be used.

Historical biogeography explains how historical geological, climatic and ecological conditions affect the existing distribution of species. Historical biogeography involves evaluation of the past geographical and evolutionary relationships of organisms. Two key developments during the 1960s were mainly important in laying the groundwork for modern phylogeography; the first was the spread of cladistic thought, and the second was the development of plate tectonics theory (De Queiroz 2005). Recent approaches, integrating coalescent theory or genealogical history of the allele and hence the distributional information can be more accurately address the relative roles of these different historical forces in determining current pattern (Cruzan and Templeton 2000). It helps in inferring the past events such as population expansion, population bottlenecks and migration. Over all, it resulted in to thinking

of vicariance biogeography, which explained the source of new descendant through geological events such as drifting apart of continents and formation of rivers. When a continuous population (or species) is divided by a new river or a new mountain range (vicariance event), two populations (or species) are created. Paleogeography, geology and paleoecology are all important fields that deliver information that is integrated into phylogeographic analysis.

Phylogeography takes a population genetics and phylogenetic perspective on biogeography. In the mid 1970s, population genetic analysis turned to mitochondrial markers (Avise 1998). The discovery of the polymerase chain reaction (PCR), a process where millions of copies of a DNA segment can be replicated, was vital in the progress of phylogeography. This breakthrough was helpful in the access of information contained in mitochondrial DNA sequences. Progress in the capillary DNA sequencing technology allowed easier sequencing of DNA and computational methods which assisted in better use of the data e.g. employing coalescent theory. It helped in improved phylogeographic inference (Avise 1998).

Early phylogeographic work has recently been criticized for its narrative nature and lack of statistical severity because it did not statistically test alternative hypotheses. Alan Templeton's Nested Clade Analysis was the only real method which made use of a conjecture key to determine the validity of a given process for explaining the concordance between geographic distance and genetic relatedness. Recent approaches have taken a stronger statistical method to phylogeography than it was done initially (Knowles and Maddison 2002; Templeton *et al* 1995; Templeton 1998).

The most favorably used molecular marker for the phylogeographic studies has been animal mitochondrial DNA (mtDNA). The mtDNA act as a bridge between systematics and population genetics. After the beginning of phylogeography (Avise *et al* 1994), it has been used mainly in examining geographical structuring of gene lineages within a species. Normally individuals are sampled from throughout the geographical range of a species. The resulting haplotypes are subsequently used to infer a phylogeny or gene tree, which reflects the evolutionary relationships of the individuals and population. By combining the consequential gene tree with the geographical location from where each individual was sampled, one can explain the

geographical distribution of major gene lineages that comprises the gene tree. Genetic structure of the sampled populations can be evaluated using of mtDNA gene sequence based genealogy and the geographic information.

Studies on population genetic structure provide scope to enlighten the fact that the fundamental evolutionary forces of selection, gene flow and genetic drift play an important role in processes such as local adaptation and speciation (Barton *et al* 1990). Recent experimental and theoretical advances have led to the increased availability of the variety of molecular markers and new methods for analyzing data using information on heterozygosity, allelic diversity and distributions. It also raises the hopes of collecting more inclusive pictures of genetic structure and deeper impending into the breeding structure. These advances have led to the causes and consequences of genetic variation among individuals of a species at population level (Avise 1994; Mitton 1994; Roderick 1996).

Evaluating the genetic variation with both individuals and populations is essential to most of the population level studies. The key factor in determining population viability is its effective population size (Frankham 1995a). It may not be necessary that individuals of a population are reproducing; therefore, the estimated effective population size is often much lower than the number of individuals alive. It is entirely based on the number of breeding individuals. This may be due to an imbalanced sex ratio, variation among individuals in number of offspring produced, causing large variation in population over time. Alternatively, it can be explained when the effective population size of a population is reduced, inbreeding is increased and causing reduction in genetic variation of a species and resulting into its capacity to adapt changing environments (Carson and Tempton, 1984; Barret and Kohn, 1991). This occurrence is generally referred as a genetic bottleneck. This is a condition when a population is greatly reduced in size and resulting into loss of rare alleles in the population. This is due to no individuals are alive who possess those rare alleles. The overall fitness of the individuals declines due to reduction in the heterozygosity level because of lack of alleles present in the population (Frankham 1995b).

The recent past has also witnessed a rapid development in the area of molecular techniques. Molecular genetics potentially helps in making faster and well-informed

decisions in conservation (O'Brien 1994). Recent advancement in DNA technology and use of molecular markers allowed a higher accuracy to investigations. DNA technology can now potentially be used to identify parents, offspring and close relatives in a single group or population and to quantify the genetic variability of present and past populations. It can also be used for reconstructing the phylogenetic relationships of very rare or extinct taxa, and in matching sample of individuals to each other and to species or populations for forensic purpose. The quantity of material required for DNA analysis can be a minute such as single hair, serum, archaeological or museum samples of pelts and bones.

Current trend is the use of non-invasive sampling (i.e. without a noticeable impact on the population during the study) of wildlife population to obtain genetic information (Morin *et al.* 1994). Genetic analysis using non-invasive sampling allows indirect measure of important population parameters (Constable *et al* 2001). It also provides an opportunity to examine, if genetic erosion may be a decisive factor for species recovery.

1.6.2 Effect of landscape and other barriers in dispersal of cervids

In highly mobile species that are distributed across continuous habitat, persistent gene flow can throttle genetic differentiation and speciation. Effective dispersal for a constant gene flow is one of the key evolutionary processes influencing population genetic structure. Naturally populations are normally subdivided into separate subpopulations connected by different degrees of gene flow. The use of population genetic studies to understand how dispersal affects genetic variation within and between populations through gene flow can not only provide new insights about the evolutionary biology and ecology of species but it can also provide precious information for conservation and management policies (Crandall *et al* 2000; Moritz 2002). However, gene flow between populations is an intricate process that is governed by many intrinsic factors such as the native dispersal ability and breeding system of a species and also by extrinsic factors such as landscape features or other environmental factors.

The affect of landscape features on dispersal and finally on population genetic structure has long been recognised. However, recently with the emergence of the new

discipline of landscape genetics, which combines landscape ecology and population genetics, exhibited the effect of landscape features on genetic population structure (Manel *et al* 2003). Though putative gene flow barriers can be exposed by associating the presence of certain landscape features with observed genetic discontinuities, it is often difficult to separate which particular features are the most likely to be affecting dispersal due to strong association between individual landscape features.

The incorporation of geographical information systems and statistics into landscape genetic studies is providing powerful tools for visualising and quantifying the consequence of landscape features on population structure in an extensive variety of organisms. For medium to large sized mammals, studies on the effect of landscape features on population genetic structure have largely been assessed for carnivores and have focused at very large geographical scales. For mammals with much narrower ranges than carnivores, landscape genetic studies are rare. In ungulates, landscape features behaving as gene flow barriers, which has been indicated or indirectly inferred from observed spatial population genetic structure in red deer, *Cervus elaphus* (Hartl *et al* 2005) and roe deer, *Capreolus capreolus* (Coulon *et al* 2006a, b). However, the influence of landscape features on population structure has only rarely been assessed using a direct quantitative method (Coulon *et al* 2004; Epps *et al* 2005). Using a model that evaluated the extent of forest habitat in a study area in southern France, it was found that female roe deer gene flow was drastically linked to forested areas (Coulon *et al* 2004).

The effects of anthropogenic barriers in isolation-by-distance regressions indicated that highways, canals and developed areas less than 40 years old were causing a rapid decrease of gene flow between desert bighorn sheep (*Ovis canadensis nelsoni*) populations in southern California (Epps *et al* 2005). The red deer is one of the largest terrestrial mammals in Britain and is widely distributed across Scotland, with the highest densities being found in the central and eastern highlands (Clutton-Brock and Albon 1989). Red deer are known for a highly mobile animal and capable of not only dispersing great distances but also crossing many kinds of terrain, including swimming across large water bodies. However, it is to be expected that some sorts of terrain are more costly to cross than others. Hence, individuals might not achieve maximum dispersal distances and, therefore, lower dispersal rates might occur

between some geographical areas (Sugg *et al* 1996). The mark–recapture data from different red deer calf-tagging programmes (these tags were recovered from dead or shot individuals), carried out at different periods of times between 1967 and 2001 on the Scottish mainland. It showed that although some individuals disperse over large distances (> 50 km), mean overall dispersal distances are commonly small, 3.3–7.4 km for males and 1.9–3.5 km for females (Daniels and McClean 2003).

The lakes, rivers and railways were identified as landscape features facilitating red deer gene flow in the Scottish highlands. In the study area, inland lakes are long but narrow and rivers are < 100 m wide, and therefore, the cost of swimming lochs and rivers might be smaller than walking around them (Pérez-Espona *et al* 2008). Red deer swimming across inland lakes is a frequently observed behaviour in the field. Lakes and rivers have been found to be dispersal barriers for other mammals with high vagility, the dimensions in the Great Lakes in Wisconsin, USA, for wolf *Canis lupus* (Mladenoff *et al* 1995); in the Peace River in western Canada for reindeer *Rangifer tarandus* (McLoughlin *et al* 2004). Lakes and rivers were not detected as a barrier for wolf dispersal in northern Poland (Jedrzejewski *et al* 2004). White *et al* (2000) observed that although bear dispersal decreased across the Mississippi river (1600 m wide), gene flow did not decrease across the White river (200 m wide). Railways seem to not have a dominating effect on the red deer population genetic differentiation as similar amount of the genetic differentiation was explained by distance alone. Nevertheless, mountain slopes, roads and forests have also been recognized as red deer gene flow barriers (Pérez-Espona *et al* 2008).

The preference of deer to move along valleys rather than across mountains can be described by the fact that upslope movement for deer is more energetically costly than horizontal movement (Parker *et al.* 1984). The effect of landscape on mammal movement has been well examined (Alexander and Waters 2000). A study in Banff National Park (Canada) showed that optimal movement of wapiti, marten (*Martes americana*), wolf, bobcat and cougar (*Puma concolor*) was through low topographic complexity areas and slopes lower than 5° (Alexander and Waters 2000). In the Scottish highlands, a similar trend was also observed, and populations sampled in the central part of the study area where terrain is less steep showed much less genetic differentiation than between populations sampled close to mountainous areas.

Main roads were also recognized as gene flow barriers to red deer. The road with an average of about 7000 cars/day was the major tourist and trade route in Scotland. It could potentially be responsible for the differentiation of the most easterly sampled estates (Pérez-Espona *et al* 2008). The main road in with the average about 4000 cars/day also probably affects the movement of red deer across the Great Glen. Negative effects of roads and railways on dispersal have also been reported for other mammals with high dispersal capabilities (Jedrzejewski *et al* 2004; Epps *et al* 2005; Waller and Servheen 2005; Riley *et al* 2006). The effect of roads on red deer gene flow needs to be explained in combination with the effect of other landscape features such as steep mountain slopes. Roads built along flatter areas corresponding to valleys and steeper terrain on either side of roads have probably been a hindrance to the movement of red deer as deer preferring to move along flatter areas rather than across them (Pérez-Espona *et al* 2008).

Forested areas have been found to influence connectivity between populations of other deer species including roe deer, white-tailed deer and red deer gene flow (Coulon *et al* 2004, Long *et al* 2005; Pérez-Espona *et al* 2008). It might be described by the fact that many forests in Scotland are to some extent fenced in order to avoid woodland damage by deer (Clutton-Brock and Albon 1989), and thus, restrict the movement of red deer (Pérez-Espona *et al* 2008). Fenced forested areas in Spain have also affected the gene flow in red deer (Martínez *et al* 2002). Other than the effect of fences on deer dispersal, deer culling regimes in forests are much higher than in open hill areas in order to keep densities low; and therefore, even if fences are crossed, the risk of mortality is high (Ratcliffe 1984).

The temporal aspect of the effect of landscape features on red deer gene flow is also equally important. Due to the longer prevalence in the landscape, the effect of natural geographical features such as sea lochs and mountain slopes in population structure is appeared to be detected than the effect of more recent human-made landscape features. However, human-made features such as roads and forests described a substantial amount of the genetic differentiation despite that its effect might be underestimated (Pérez-Espona *et al* 2008). The very old roads during the end of the 19th century and erected fences along the roads would have hindered the movement of deer (Pérez-Espona *et al* 2008). The same applies for fenced forests in Scotland,

where it was heavily deforested by the end of the 17th century and during few decades ago the large programmes of reforestation at commercial scale was took place (Clutton-Brock and Albon 1989). Assuming that the effect of these two human-made features (landscape barrier and temporal effect) is not only due to adjacent natural features and considering an average generation time for red deer of about 8 years (Kruuk *et al* 2002), main roads and fenced forests could potentially be considered important contemporary red deer gene flow barriers as their effect could be detectable in less than 10 generations (Pérez-Espona *et al* 2008).

It is also significant to consider that the intrinsic correlation between landscape variables might confound the effect of a particular landscape feature on population genetic structure. The correlation between landscape features has been assessed in several studies (Roach *et al* 2001; Spear *et al* 2005; Cushman *et al* 2006), these have generally involved partial Mantel tests (Raufaste and Rousset 2001; Rousset 2002). The redundancy distance-based analyses have been proven significant in assessing the interaction between different ecological variables in landscape ecology studies (McArdle and Anderson 2001). It could be considered as an alternative to partial Mantel tests in landscape genetics studies. The development of analytical methods in landscape genetics still remains a challenge (Storfer *et al* 2007).

1.6.3 Effect of refugia camp in evolutionary process

Phylogeographical studies in South America have recognized multiple refugia (Turchetto-Zolet *et al.* 2013). The contrasting pattern, demographic expansion from a single refuge was identified largely in mammals from temperate South America (Palma et al. 2005; Rodriguez-Serrano *et al*. 2006; Gonzalez-Ittig *et al*. 2010; Lessa et al. 2010). The existence of some species in multiple refugia confined to a small area throughout their distribution range indicates that these species probably have persisted through multiple climatic cycles in heterogeneous environments. It highlighted the importance of dynamic evolutionary processes and a mosaic of habitats in heterogeneous landscapes that allowed species to persist through changing environmental conditions.

Recent palaeoecological information in the Neotropics indicates that during the quaternary, spatial reorganization and persistence in appropriate microrefugia were

more frequent than absolute extinctions (Vegas-Vilarrubia *et al.* 2011). Microrefugia are known as small areas with local favourable environmental features in which a small population can survive outside their main distribution area protected them from unfavourable regional environmental conditions (Rull 2009). Pollen records suggest that in the southern Andes, ice-free areas might have assisted in the maintenance of habitat refugia for forest taxa (Markgraf *et al.* 1995; Heusser 2010).

1.6.4 Biogeography of India

The peninsular India is geomorphologically quite distinct and biographically from the rest of India and its primary faunistic affinities are to be tracked back more to the Madagascan Region than to the Oriental or even the Malayan area. Considering this point, three principal amphitheatres of origins of floras and faunas, diffentiations and radiations have contributed largely to the compositions and evolution of the biogeography of India. These amphitheaters are (a) the Peninsula, (b) the higher Himalaya and (c) the eastern amphitheatres of the Indo-Chinese and Malayan sub regions of the Oriental Realm as represented in figure 1.2 (Mani 1974).

The floras and faunas differentiated in these amphitheatres fall respectively into two major groups, the representatives of which constitute the biogeographical components today. These two groups are (a) the Gondwana derivatives and (b) the Asiatic derivatives. The fauna that differentiated in the Eastern amphitheatres is similar to a humid tropical forest one. However, unlike the Peninsular faunas is mainly composed of phylogenetically much younger and taxonomically higher groups derives almost exclusively from Asia. These faunas can be characterized by a high of phylogenetic plasticity, evolutionary intensity and ecologically and geographically volatile (Mani 1974). In noticeable contrast to the peninsular faunas, the eastern faunas are not older than perhaps the Pliocene, Pleistocene and even Post Pleistocene times (Mani 1974). Their evolution and dispersal were conditioned mainly by and closed bound up with the immense tertiary mountain-building activities in these areas. These faunas are characteristically rich in snakes, higher mammals, birds, forest and arboreal insect like Orethopera, Phasmida, Cerambycidae and Lepidoptera. The faunas spread west wards along the forest-covered ranges of the Himalayas and to the Peninsula can be called as replacing faunas in the Peninsula, the original fauna of which is of much older age (Mani 1974).

(a) The Peninsular amphitheatres, which is an ancient endemic relict fauna;

(b) The Indo-Chinese amphitheatre, largely outside limits of India, but extending west wards as a narrow belt on the forest covered ranges of the Himalaya, and overlapping the peninsular amphitheatres in the northeast Assam;

(c) The Malayan amphitheatres overlapping the Indo-Chinese amphitheatre in the north;

(d) The Turkmenian amphitheater, also outside limits of India, except for a narrow southern fringe encroaching on the higher Himalaya above the forest line and an area of differentiation of young steppes fauna.

The amphitheatre b and c are regions of differentiation of young, tropical-forest faunas. The amphitheatres c-d represent the biogeographical appendages of India (Figure 1.2).

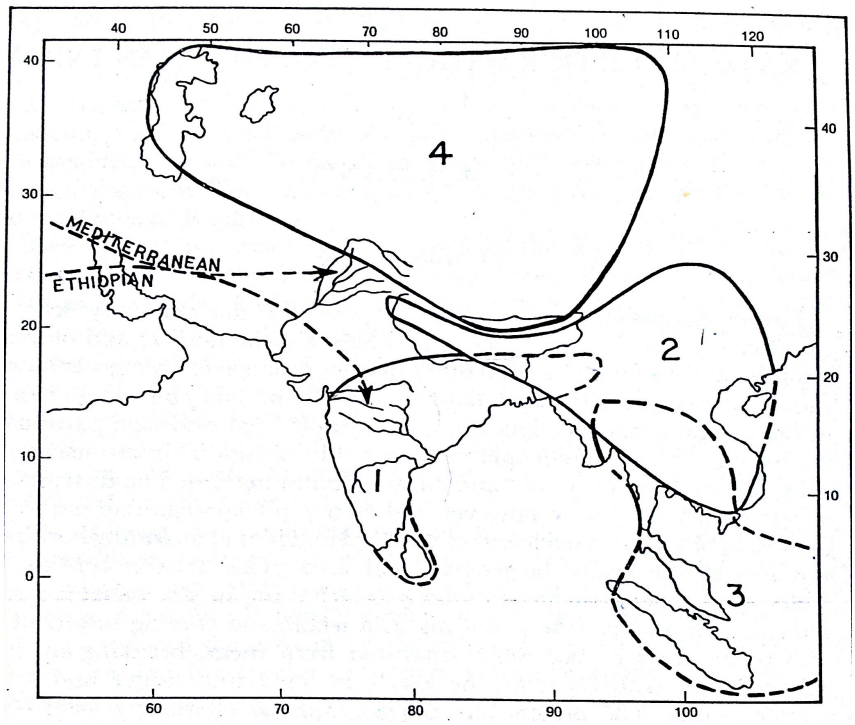


Figure 1.2: Major amphitheatres of differentiation, evolution and radiation of floras and faunas, which have profoundly influenced the biogeographically evolution of India (Source: Mani 1974).

1.6.5 Conservation genetics

In context to the above introduction, the proposed research work will be entirely based on the methodology of molecular genetics i.e., the use of molecular markers. The molecular markers are supposed to play a key role in inferring the genetic analysis of sambar at individual as well as the population level across the world and within the Indian subcontinent. Analysis of fecal DNA extends a great support to the research in the area of conservation genetics (Piggot *et al* 2003, 2004), which allows the maximum use of field samples without harming the animals.

Conservation genetics or the application of genetics to the preservation of species has received increasing attention in recent years (Allendorf and Luikart, 2007; Frankham, 2003). In conservation genetics, knowledge of the relatedness between individuals is particularly important in captive breeding programs that seek to reduce incestuous mating in order to minimize inbreeding and the loss of genetic variation (Frankham *et al* 2002). It is well established that a decline in genetic variation reduces the ability of a population to adapt to environmental changes and therefore decreases its long term survival. The loss of genetic diversity also results in lower individual fitness and poor adaptability (Lande, 1988). The fate of small populations is linked to genetic changes. The captive breeding of endangered wildlife animals is often necessary for their conservation; however, this strategy potentially increases the chances of inbreeding that, in turn, causes poor fitness of these populations (Ralls and Ballou, 1983; Crnokrak and Roff, 1999). Inbreeding is known to decrease genetic diversity and to reduce reproductive and survival rates leading to increased extinction risk. Genetically impoverished endangered populations often fail to exhibit signs of recovery until crossed with individuals from other populations (Land and Lacy 2000). Moreover, wildlife populations with lower degree of genetic diversity are at greater risk of extinction (Saccheri *et al* 1998). Knowledge and studies on genetics can reduce the extinction risk by helping to develop appropriate population management programs that can minimize the risks implied through inbreeding. Breeding programs are often started assuming that the wild founders initiating the captive population are unrelated. However, threatened animals brought into captivity often have small population sizes and therefore the founders may be related to each other (Geyer *et al* 1993; Haig *et al* 1994). It is well-known that decline in genetic variation reduces the ability of a population to acclimatize to environmental changes and hence decreases the chances

for its long term survival (Arif and Khan 2009). Evaluation and conservation of biodiversity of wild populations is critically important to minimize the loss of initial genetic variation as a corollary of inbreeding (Russello and Amato, 2004). Molecular methods play an important role in estimating the genetic diversity among individuals by comparing the genotypes at a number of polymorphic loci (Avise, 2004). Several types of molecular markers, including mitochondrial DNA (mtDNA) and nuclear DNA markers, are available but none of them can be regarded as optimal for all applications (Sunnucks, 2000).

1.6.6 Molecular markers

Molecular markers are indispensable tools for determining the genetic variation and biodiversity with high levels of accuracy and reproducibility. These markers are mainly classified into two types; mitochondrial and nuclear markers. The widely used mitochondrial DNA markers with decreasing order of conserved sequences are 12S rDNA > 16S rDNA > cytochrome *b* > control region (CR); thus the 12S rDNA is highly conserved and the CR is highly variable. The most commonly used nuclear markers for DNA fingerprinting include Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and Short Tandem Repeats (STR's) or Microsatellites.

1.6.6.1 Mitochondrial DNA

Mitochondria are structures within cells that convert the energy in the form of adenosine triphosphate (ATP) from food into a form that cells can use. Although most DNA is packaged in chromosomes within the nucleus, mitochondria also have a small amount of their own DNA. In humans, mitochondrial DNA can be assessed as the smallest chromosome coding for only 37 genes and containing only about 16,600 base pairs (Figure 1.3). Human mitochondrial DNA was the first significant part of the human genome to be sequenced. In most species, including humans, mtDNA is inherited solely from the mother (Bradshaw Foundation 2012). The DNA sequence of mtDNA has been determined from a large number of organisms and individuals (including some organisms that are extinct), and the comparison of those DNA sequences represents a mainstay of phylogenetics, in that it allows biologists to elucidate the evolutionary relationships among species. It also permits an examination

of the relatedness of populations, and so has become important in anthropology and field biology.

Nuclear and mitochondrial DNA are thought to be of separate evolutionary origin, with the mtDNA being derived from the circular genomes of the bacteria that were engulfed by the early ancestors of today's eukaryotic cells. This theory is called the endosymbiotic theory. Each mitochondrion is estimated to contain 2-10 mtDNA copies (Wiesner *et al* 1992). In the cells of extant organisms, the vast majority of the proteins present in the mitochondria (numbering approximately 1500 different types in mammals) are coded for by nuclear DNA, but the genes for some of them, if not most, are thought to have originally been of bacterial origin, having since been transferred to the eukaryotic nucleus during evolution.

In most multicellular organisms, mtDNA is inherited from the mother (maternally inherited). Mechanisms for this include simple dilution (an egg contains 100,000 to 1,000,000 mtDNA molecules, whereas a sperm contains only 100 to 1000), degradation of sperm mtDNA in the fertilized egg, and, at least in a few organisms, failure of sperm mtDNA to enter the egg. Whatever the mechanism, this single parent (uniparental) pattern of mtDNA inheritance is found in most animals, most plants and in fungi as well.

The fact that mitochondrial DNA is maternally inherited enables genealogical researchers to trace maternal lineage (Y-chromosomal DNA, paternally inherited, is used in an analogous way to trace the agnate lineage). This is accomplished on human mitochondrial DNA by sequencing one or more of the hypervariable control regions (HVR1 or HVR2) of the mitochondrial DNA, as with a genealogical DNA test. HVR1 consists of about 440 base pairs. These 440 base pairs are then compared to the control regions of other individuals (either specific people or subjects in a database) to determine maternal lineage. Most often, the comparison is made to the revised Cambridge Reference Sequence. Vilà *et al* (1997) have published studies tracing the matrilineal descent of domestic dogs to wolves. The concept of the mitochondrial eve is based on the same type of analysis, attempting to discover the origin of humanity by tracking the lineage back in time as mtDNA is not highly conserved and has a rapid mutation rate; it is useful for studying the evolutionary

relationships phylogeny of organisms. Biologists can determine and then compare mtDNA sequences among different species and use the comparisons to build an evolutionary tree for the species examined

Evidence supports rare instances of male mitochondrial inheritance in some mammals as well. Specifically, documented occurrences exist for mice (Gyllensten *et al* 1991; Shitara *et al* 1998), where the male-inherited mitochondria were subsequently rejected. It has also been found in sheep (Zhao *et al* 2004), and in cloned cattle (Steinborn *et al* 1998). It has been found in a single case in a human male (Schwartz and Vissing 2002).

The mtDNA contains 37 genes, all of which are essential for normal mitochondrial function. Thirteen of these genes provide instructions for making enzymes involved in oxidative phosphorylation. Oxidative phosphorylation is a process that uses oxygen and simple sugars to create adenosine triphosphate (ATP), the cell's main energy source. The remaining genes provide instructions for making molecules called transfer RNA (tRNA) and ribosomal RNA (rRNA), which are chemical cousins of DNA. These types of RNA help assemble protein building blocks (amino acids) into functioning proteins.

1.6.6.1.1 Mitochondrial ribosomal RNA gene

Animal mitochondria contain two ribosomal RNA (rRNA) genes, 12S rDNA and 16S rDNA. Mitochondrial 12s rDNA is highly conserved and has been applied to understand the genetic diversity of higher categorical levels such as in phyla. On the other hand, the 16s rDNA is often used for studies at middle categorical levels such as in families or genera (Gerber *et al* 2001). For molecular analysis, these markers are first amplified by PCR using conserved primers and the amplicons are sequenced. Sequencing data are then aligned and compared using appropriate bioinformatics tools. Alvarez *et al* (2000) have suggested specific haplotypes of 12S rRNA gene to study the effects of geographical isolation on genetic divergence of endangered spur–thighed tortoise *(Testudo graeca*). A short fragment of 12S rRNA gene was used to examine genetic variation in 158 tortoise specimens belonging to the four different subspecies of *Testudo graeca* (Van der Kuyl *et al* 2005).

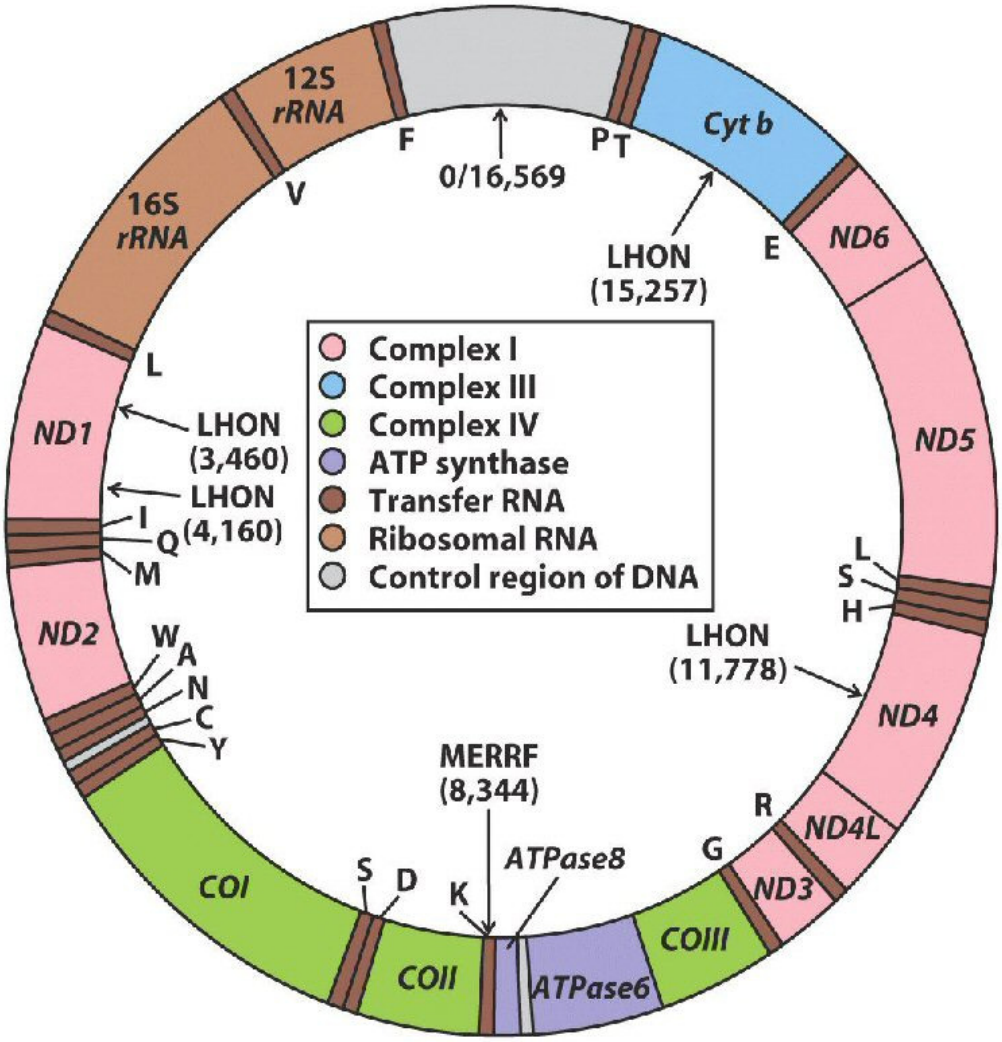


Figure 1.3: Diagrammatic representation of a Mitochondrial DNA

The mitochondrial DNA haplotyping has suggested that the tortoise subspecies of *Testudo graeca graeca* and *Testudo graeca ibera* are genetically distinct with a calculated divergence time in the early or middle Pleistocene; however, other proposed subspecies could not be recognized based upon their mitochondrial haplotypes (Van der Kuyl *et al* 2005). The 12S rRNA fragment of *Testudo graeca* was found to be somewhat less variable than the D–loop fragment, due to the inherently slower evolutionary rate of rRNA genes than the variable parts of the D–loop (Pesole *et al* 1999). Lei *et al* (2003) examined the mitochondrial rRNA genes of Chinese antelopes and observed that average sequence divergence values for 16S and 12S rRNA genes were 9.9% and 6.3% respectively. A single base in the 16S rDNA sequences from the endangered species *Pinna nobilis* was found to be different in all analyzed individuals from a single population sample differentiating it from others (Katsares *et al* 2008).

Mitochondrial 16S rRNA was used to elucidate the pattern of relationships and systematic status of 4 genera, including nine species of skates living in the Mediterranean and Black Seas (Turan, 2008). Molecular studies on endangered Pecoran have shown lower sequence diversity in 16S rRNA gene as compared to cytochrome b gene, both between and within species; however, the 16S rRNA gene harbored a larger number of species–specific mutation sites than cytochrome b gene (Guha *et al* 2006). Na-Nakorn *et al* (2006) have assessed the level of genetic diversity in critically endangered Mekong giant catfish species using 570 bp sequences of 16S rRNA from 672 individuals of nine species. In all species studied, haplotype diversity and nucleotide diversity ranged from 0.118–0.667 and 0.0002–0.0016, respectively. Four haplotypes were detected among 16 samples from natural populations of Mekong giant catfish. The findings from this study have important implications for conservation of the Mekong giant catfish, especially in designing and implementing artificial breeding program for restocking purposes (Na-Nakorn *et al* 2006)

1.6.6.1.2 Mitochondrial protein–coding genes

Due to their faster evolutionary rates compared to ribosomal RNA genes, the mitochondrial protein–coding genes are regarded as powerful markers for genetic diversity analysis at lower categorical levels, including families, genera and species.

Animal mitochondria contain 13 protein–coding genes; however, one of the most extensively used protein coding genes of the mitochondrial genome for molecular analysis is cytochrome *b* (cyt *b*). Mitochondrial cyt *b* sequences have been used to understand the genetic diversity for better conservation management of Tibetan gazelle (*Procapra picticaudata*), a threatened species on the Qinghai–Tibet Plateau of China (Zhang and Jiang, 2006). The sequence analysis of 46 samples collected from 12 geographic locations identified 16 cyt *b* haplotypes, to be used as molecular markers for conservation planning (Zhang and Jiang, 2006). Partial cyt *b* based molecular analysis of genetic distances has revealed that there is considerable genetic divergence between the Korean goral and the Chinese goral, but virtually none between Korean and Russian gorals (Min *et al* 2004). The Korean gorals possessed two haplotypes with only one nucleotide difference between them, while the Japanese serows showed slightly higher sequence diversity with five haplotypes. These data highlight the importance of conservation of the goral populations of these regions, and the need to reconsider the taxonomic status of Korean and Russian gorals (Min *et al* 2004). Another important mitochondrial protein coding gene, NADH dehydrogenase subunit 5 (318 bp), has been used for phylogenetic analysis of multiple individuals of different species from Felidae family and successfully differentiated eight clades reflecting separate monophyletic evolutionary radiations (Johnson and O'Brien, 1997). Mitochondrial cytochrome oxidase I (COI) gene has recently gained more attention in developing DNA barcodes for species identification and biodiversity analysis.

1.6.6.1.3 Mitochondrial control region

Mitochondrial DNA contains a non–coding region termed the control region (CR or D–loop) due to its role in replication and transcription of mtDNA. The D–loop segment exhibits a comparatively higher level of variation than protein–coding sequences due to reduced functional constraints and relaxed selection pressure. The length of the D–loop is approximately 1 kb and it can easily be amplified by PCR prior to sequencing to determine the molecular diversity. Sequence analysis of the CR of the sun bear has been used to measure molecular diversity and to identify conservation units for better management of this species (Onuma *et al* 2006). Wu *et al* (2006) have sequenced a portion of mitochondrial CR (424 bp) to assess the population structure and gene flow among the populations of black muntjac

(*Muntiacus crinifrons*) using 47 samples collected from three large populations. A total of 18 unique haplotypes (15 of them as population–specific) were defined based on 22 polymorphic sites.

It has been suggested that the coexistence of distinct haplotypes in a specific population was induced by historical population expansion after fragmentation and that the current genetic differentiation should be attributed to the reduction of female–mediated gene flow due to recent habitat fragmentation and subsequent loss (Wu *et al* 2006). Hu *et al* (2006) have studied the genetic diversity and population structure of the Chinese water deer (*Hydropotes inermis inermis*) by analyzing the 403 bp fragment of mitochondrial D–loop. They have detected 18 different haplotypes in 40 samples demonstrating the haplotype diversity of 0.923 and nucleotide diversity of 1.318, whereas no obvious phylogenetic structure among haplotypes was found for samples of different origin (Hu *et al* 2006). Iyengar *et al* (2006) have performed a comparative study of CR sequences from several captive oryx species and concluded a close grouping of *Oryx leucoryx* with *Oryx gazelle* instead of *Oryx dammah*. Idaghdour *et al* (2004) have sequenced the 854 bp of the CR from the houbara bustard (*Chlamydotis undulate*) to describe the molecular diversity of this threatened cryptic desert bird, whose range extends from North Africa to Central Asia. Zhang and Jiang (2006) have used CR sequences to investigate the genetic diversity and evolutionary history of the Tibetan gazelle. A total of 25 CR haplotypes with high frequencies of both CR haplotype and nucleotide diversities were identified. These findings have suggested that the present population structure is the result of habitat fragmentation during the recent glacial period on the Qinghai–Tibet Plateau and it is likely that the present populations of Tibetan gazelle exhibit a pattern reminiscent of several bottlenecks and expansions in the recent past (Zhang and Jiang, 2006).

1.6.6.1.4 Summary of the characteristics of mtDNA marker(s)

mtDNA	- Inherited from the mother (maternal lineage); rare exceptions do exist. - Evolves about 10–fold faster than nuclear DNA. - It can be used from degraded or old samples.
12s rDNA	- Highly conserved; used for high–category levels: phyla and subphyla

16S rDNA	- Usually used in mid–category differentiation such as families
Protein–coding genes	- Conserved; used in low–categories such as families, genera and species (phylogenetics)
Control region	- Highly variable region of mtDNA and used in population genetics

1.6.7 Microsatellites

Microsatellites, also known as simple sequence repeats (SSR) or short tandem repeats (STR), are non-coding repetitive DNA regions composed of small motifs of 1 to 6 nucleotides repeated in tandem, which are widespread in both eukaryotic and prokaryotic genomes (Field and Wills, 1998; Tóth *et al* 2000). It is a type of Variable Number Tandem Repeat (VNTR). Microsatellites are typically co-dominant. They are used as molecular markers in STR analysis, for kinship, population and other studies. They can also be used for studies of gene duplication or deletion, marker assisted selection, and fingerprinting. The variability of microsatellites is due to a higher rate of mutation compared to other neutral regions of DNA. These high rates of mutation can be explained most frequently by slipped strand mispairing (slippage) during DNA replication on a single DNA strand. Mutation may also occur during recombination during meiosis (Blouin *et al* 1996), although genomic microsatellite distributions are associated with sites of recombination most probably as a consequence of repetitive sequences being involved in recombination rather than being a consequence of it (Huang *et al* 2002). Some errors in slippage are rectified by proofreading mechanisms within the nucleus, but some mutations can escape repair. The size of the repeat unit, the number of repeats and the presence of variant repeats are all factors, as well as the frequency of transcription in the area of the DNA repeat. Interruption of microsatellites, perhaps due to mutation, can result in reduced polymorphism. However, this same mechanism can occasionally lead to incorrect amplification of microsatellites; if slippage occurs early on during PCR, microsatellites of incorrect lengths can be amplified.

Microsatellites can be amplified for identification by the polymerase chain reaction (PCR) process, using the unique sequences of flanking regions as primers. DNA is repeatedly denatured at a high temperature to separate the double strand, and

then cooled to allow annealing of primers and the extension of nucleotide sequences through the microsatellite. This process results in production of enough DNA to be visible on agarose or polyacrylamide gels; only small amounts of DNA are needed for amplification because in this way thermocycling creates an exponential increase in the replicated segment (Griffiths *et al* 1996). With the abundance of PCR technology, primers that flank microsatellite loci are simple and quick to use, but the development of correctly functioning primers is often a tedious and costly process.

Microsatellite analysis is a relatively new technology in the field of forensics, having come into popularity in the mid-to-late 1990s. It is used for the genetic fingerprinting of individuals. The microsatellites in use today for forensic analysis are all tetra- or penta-nucleotide repeats (4 or 5 repeated nucleotides), as these give a high degree of error-free data while being robust enough to survive degradation in non-ideal conditions. Shorter repeat sequences tend to suffer from artefacts such as PCR stutter and preferential amplification, as well as the fact that several genetic diseases are associated with tri-nucleotide repeats such as Huntington's disease. Longer repeat sequences will suffer more highly from environmental degradation and do not amplify by PCR as well as shorter sequences.

The analysis is performed by extracting nuclear DNA from the cells of a forensic sample of interest, then amplifying specific polymorphic regions of the extracted DNA by means of the polymerase chain reaction. Once these sequences have been amplified, they are resolved either through gel electrophoresis or capillary electrophoresis, which will allow the analyst to determine how many repeats of the microsatellites sequence in question there are. If the DNA was resolved by gel electrophoresis, the DNA can be visualized either by silver staining (low sensitivity, safe, inexpensive), or an intercalating dyesuch as ethidium bromide (fairly sensitive, moderate health risks, inexpensive), or as most modern forensics labs use, fluorescent dyes (highly sensitive, safe, expensive). Instruments built to resolve microsatellite fragments by capillary electrophoresis also use fluorescent dyes to great effect. It is also used to follow up bone marrow transplant patients (Antin *et al* 2001). In the United States, 13 core microsatellite loci have been decided upon to be the basis by which an individual genetic profile can be generated (Butler 2005). These profiles are stored on a local, state and national level in DNA databanks such as CODIS.

Broadly used as genetic markers, microsatellites have a particular attribute in that they suffer higher rates of mutation than the rest of the genome (Jarne and Lagoda, 1996). Microsatellites are classified according to the type of repeat sequence as perfect, imperfect, interrupted or composite. In a perfect microsatellite the repeat sequence is not interrupted by any base not belonging to the motif (*e.g.* TATATATATATATATA) while in an imperfect microsatellite there is a pair of bases between the repeated motifs that does not match the motif sequence (*e.g.* TATATATA*CT*ATATA). In the case of an interrupted microsatellite there is a small sequence within the repeated sequence that does not match the motif sequence (*e.g.* TATATA*CGTG*TATATATA) while in a composite microsatellite the sequence contains two adjacent distinctive sequence-repeats (*e.g.* TATATATATAGTGTGTG).

Although originally designed for research in humans, microsatellite analysis has become a powerful tool for research on animals (Schlötterer *et al.* 1991) and plants (Dayanandan *et al.* 1997; Cipriani *et al.* 1999; Collevatti *et al.* 2001, Roa *et al.* 2000; Steinkellner *et al.* 1997; White and Powell, 1997). Until a few years ago, microsatellites were thought to be selectively neutral markers and not affected by selective pressures. However, it is now evident that the expansion of the number of repeats may cause human diseases. For example, Huntingtons disease is caused by increases in the length of a CAG motif repeat present in the huntingtin protein gene on human chromosome 4 (Moxon and Wills, 1999), and an increasing number of neurodegenerative disorders have been related to expanded microsatellite repeats, mainly in the tri-nucleotide class (Cummings and Zoghbi, 2000; Everett and Wood, 2004; Goldstein and Schlotterer, 1999). Quite interesting is the fact that microsatellites are preferentially associated with non-repetitive DNA in plant genomes *i.e.* they frequently occur within and near genes (Morgante *et al* 2002).

1.6.7.1 Genetic features of microsatellites

A homozygous microsatellite locus has the same number of repeats on both homologous chromosomes; whereas a heterozygous microsatellite locus has a different number of repeats for each allele *e.g.* one allele can contain 9 repeats and the other 10. However, at the same locus the population as a whole usually contains several alleles each with a different number of repeats, which means that

microsatellite markers are very useful for discriminating different individuals. Assuming that *m* is the number of alleles in a population, the maximum number of different genotypes (NDG) will be $m(m + 1)/2$ and the number of possible heterozygous genotypes (NHG) will be $m(m - 1)/2$, *e.g.* if $m = 48$, NDG = 1,176 and NHG = 1,128. The high discriminating power of microsatellites is an important characteristic which justifies their use in population genetic studies and forensic science.

1.7 EVOLUTION OF GENUS RUSA (A FOSSIL RECORD)

Cervidae is a species-rich family which evolved in Eurasian and radiated from tropical to temperate climates in the Pleistocene (Gilbert *et al.* 2006; Geist 1998; Groves 2007). It contains 26 fossil genera and 80 fossil species from China alone (Dong 1993). Paleomerycines (Amphitragulus), procervulines (Procervulus), and cervulines (Eostyloceros) of the Miocene were putative precursors of species now included in the family Cervidae, with fossil species of Cervavitus which evolve to tribe Cervini (Flerov 1952; Miyamoto *et al.* 1990). It shows recent radiation into South America during the Pliocene (Eisenberg 1987). Cervidae is a family of the Northern Hemisphere (Geist 1998). The only cervid in Africa, *Cervus elaphus barbarus* (Barbary red deer) might have been introduced from Europe along the southern Mediterranean coast before 8,000 years. Several cervid fossils of Pleistocene origin indicated the transition of herbivorous species from forest to grasslands (Gilbert *et al.* 2006).

During the early Pliocene, rusine deer were found in Europe and in late Pleistocene 3-tine antlers species were evolved, such as the Philis deer (*Cervus philisi*), which showed affinity with living Rusa (Lydekker 1898; van Bemmel 1974, Kurtén 1968; Lister 1987). Sambar is among the most ancestral of living cervids, with characteristics little changed from the late Pliocene and paralleling other Chinese pliocervines (Leslie 2011). It likely evolved in southern tropical areas (Flerov 1952), perhaps from the extinct Pleistocene forms such as *Epirusa hilzheimeri* or Eucladoceros (Leslie 2011; Geist 1998; Grubb 1990). It was suspected that *Rusa elegans* branched in the mid-Villafranchian during 2.0–2.5 million years ago, giving rise to the extinct *R. hilzheimeri*, which gave rise to *R. unicolor* and the high-elevation

specialist *Przewalskium albirostre* of the Tibetan Plateau (Di Stefano and Petronio 2002; Leslie 2009; Schaller 1998).

Five rusine fossil species i.e. *elegans, microta, stehlini, unicolor* and *yunnanensis* from the early Pleistocene were found throughout China and *R. timorensis* apparently did not appear until the late Pleistocene (Dong 1993). A fossil Rusa of middle Pleistocene origin (0.8-1.7 mya) larger than existing *R. unicolor* were recovered from the caves in northern Vietnam (Bacon *et al.* 2004). In Sichuan (China), antlers of *R. unicolor* of middle Pleistocene origin were possibly used by humans prior to fossilization (Hooijer 1951). Bones of *R. unicolor* were also found in the caves of Paleolithic origin in China (Huang *et al.* 1995; Si *et al.* 1993).

1.8 RESEARCH GAPS

Rusine deers are least studied and phylogenetically highly informative group (Timmin *et al.* 2008). Recent phylogenetic revisions within this group apparently indicated that this group is taxonomically more fragile among cervides (Timmin *et al.* 2008). It has helped in separation of *R. unicolor* subspecies in to a distinct species (*R. marianna*). As sambar distribution in India is very wide and little study has been done on evaluation of its genetic features, a dedicated phylogeographic and population genetic study was needed to understand its speciation pattern, adaptive radiation and population genetic structure. In view of this, the following objectives were set forth:

1.9 OBJECTIVES OF THE STUDY

(a) To examine the intra-species relationship among sambar populations across various biogeographic zones in India,

(b) To study the molecular systematic of sambar in order to understand its phylogenetic structure and its relationships with other cervids; and

(c) To examine the genetic diversity and gene flow pattern among sambar populations

1.10 KEY RESEARCH QUESTIONS

In broader terms this research aims to find out

(a) Is there any variation in phylogenic position of sambar across biogeographic zones of India?

(b) What is the precise phylogenic position of Indian sambar with respect to other cervids?

(c) Does the population genetic structure of sambar differ across biogeographic zones of India?

(d) What is the status of gene flow in sambar across biogeographic zones?

1.11 SIGNIFICANCE OF THE STUDY

There are seven surviving subspecies reported for sambar and very few molecular studies has been done to evaluate, identify and screen the taxonomy of different population of sambar. Recently, a subspecies of *R. unicolor* has been separated as distinct species *R. marianna* in Philippines. Considering a large distribution range of sambar in India, a grave need was felt to examine the phylogenetics (phyogeography) and level of genetic diversity in different sambar populations of India. In this study these issues were addressed. Molecular signatures and marker for differentiation of various sambar populations has been established during this study. It also provided molecular insight in to the taxonomic status and the level of genetic diversity of among sambar populations. It laid a milestone for further study of similar widely distributed species.

CHAPTER 2
MATERIAL AND METHODS

SUMMARY

This chapter describes the geographical location of collected biological sample and protocols for sample collection, laboratory analysis and record of morphological trait. The biological samples such as fecal pellets, tissue and antlers were collected from the distribution range of Sambar in India, representing various biogeographic zones for genetic analysis. The validated forensic samples of Sambar were also used in this study. The analysed samples were from Western Ghats, Deccan Plateau, Eastern Ghats and Gangetic Plains. The present chapter also describes the comparison of various DNA extraction protocols for successful extraction of PCR amplifiable quality DNA from non-invasive biological samples. We compared three different DNA extraction protocols during this study. These protocols were phenol-chloroform (PC), commercial kit and Guanidine hydrochloride (Gu-HCl) based in-house method. The effectiveness of the above protocols was compared based on comparative success rate of PCR amplification form the DNA extracted from bones, antlers and faecal samples. This study highlights that silica based DNA extraction protocol using lesser toxic Gu-HCl method yields the better quality of DNA with negligible PCR inhibitors.

2.1 STUDY AREA

India is located at the tri-junction of the biodiversity rich Ethiopian, Indo-Malayan and Palaearctic realms. It has led to high species richness in India. India can be divided into five physiographic regions. These regions are the Northern Mountains, Indo Gangetic Plains, The Peninsular Plateau, Thar Desert and The Coastal Plains. It is amongst the few countries that have developed a biogeographic classification for conservation planning, and has mapped biodiversity-rich areas in the country. Biogeographic zones are the large distinctive units of similar ecology, biome representation, community and species. Biogeographic classification of India was done by Rodgers and Panwar (1988), describing 10 biogeographic zones in India, further divided into 25 biogeographic provinces. The maps were further revised by

Rodgers, Panwar and Mathur (2002), using GIS techniques into 10 zones and 26 provinces. The classification was done using various factors such as altitude, moisture, topography, rainfall, etc.

Biogeography deals with the geographical distribution of plants and animals. Communities of plants and animals in different geographical areas of the world differ widely from each other. Biogeography is divided into branches:

i) Phyto-geography and

ii) Zoo-geography

Phyto-geography (plant geography) deals with origin, distribution and environmental interrelationships of plants. Zoogeography deals with the migration and distribution of animals. From the stand point of biodiversity, India is given the status of a "Mega-diversity" zone. The 10 biogeographic zones (Figure 2.1) of India are as follows:

Table 2.1 Coverage area of National Parks and Protected Area's in Biogeographic zones of India

Zone	Name	Geographical Area (km^2)	% as PA's
1	Trans Himalayan	184823	9.26
2	Himalayan	210673	11.89
3	Desert	214014	7.51
4	Semi-Arid	539479	2.63
5	Western Ghats	132179	11.12
6	Deccan Peninsula	1380339	4.15
7	Gangetic plain	354848	2.19
8	Coasts	91319	5.75
9	North East	171340	3.74
10	Island	8249	18.71

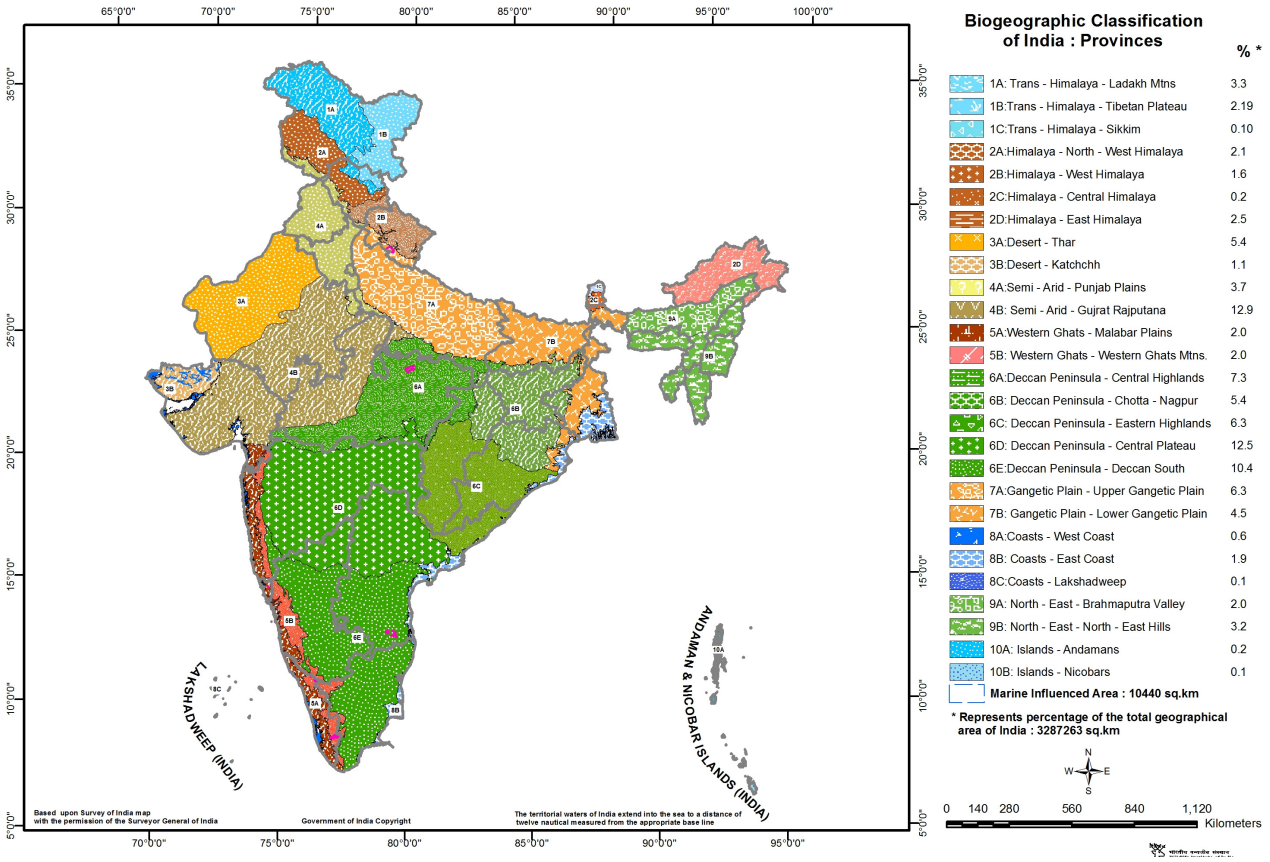


Figure 2.1: Biogeographic zones of India

(Source: WII, GIS CELL)

2.1.1 Trans Himalayan and Himalayan

This is the northern most area of the country around Himalayas. This zone is not related to mountains but it is the area present surrounding the mountains. This region shows irregular vegetation. It has the richest wild sheep producing quality wool. Snow leopard is a special animal observable in this zone. Migratory birds like black neck crane are seen here. The great Indian bustard which is an endangered species has also been reported in the grasslands west to this zone. Himalayan is a hilly region with good flora and fauna, exhibiting maximum biodiversity in this region.

2.1.2 Desert and Semi-Arid

Desert is a part of Rajasthan state, from where, the great desert of western India namely 'Thar desert' starts, and Desert zone is really a deserted zone with areas like Jaisalmir surrounded by sand dunes, which are the best centres for studying soil erosion by wind. Semi-Arid starts in Rajasthan and extends up to some parts in the state of Punjab and Haryana. In this area, ground water and surface water is much less. Dry xerophytic vegetation is predominant, and faunal diversity is minimum.

2.1.3 Western Ghats

The Western Ghats or the Sahyadri constitute a mountain range along the western side of India. It is a UNESCO World Heritage Site and is one of the eight "hottest hotspots" of biological diversity in the world. The range runs north to south along the western edge of the Deccan Plateau, and separates the plateau from a narrow coastal plain along the Arabian Sea. It represents the mountainous western zone of south peninsular India having rich flora and fauna with tropical rain forests extending from Konkan region of Maharashtra up to the western part of Kerala generally called Malabaar coast of Arabian Sea. Wild relatives of cultivated plants like banana, mango, citrus, black pepper are found abundantly in this part. At least 325 globally threatened species occur in the Western Ghats.

2.1.4 Deccan Peninsula

This is the central table of south India with rich flora and fauna. But some areas represent semiarid type of vegetation. Rainfall occurs mainly by the south west monsoon and so limited number of dense forests is present in Deccan. Deccan platean

is represented in the states of Madhya Pradesh, Maharashtra, Orissa, Andhra Pradesh, Tamilnadu, Karnataka and Kerala.

2.1.5 Gangetic Plain

Gangetic plain zone is the most important zone in terms of human concentration in India. It is a vast enclosed basin of numerous small and large rivers, separated by alluvial divides. The western section, comprising Punjab, Haryana, Chandigarh, Delhi and western Uttar Pradesh, is slightly higher in elevation. Geologically, the whole region is made up of alluvium brought down by the Himalayan rivers. The upper Ganga plain is a vast stretch of the Indo-Gangetic Plains, where the rivers play an important role and have a definite influence over the area. The Middle Ganga plain is the transitional zone between the upper Ganga plain and the Lower Ganga plain. The Lower Ganga plain extends over Bihar and West Bengal and has a higher rainfall.

2.1. 6 Coasts

Coastal India spans from the south west Indian coastline along the Arabian sea. It starts from the coastline of the Gulf of Kutch in its western most corners. It stretches across the Gulf of Khambhat, and down through Cape Comorin in the southernmost region of South India. In the eastern part it run through the Cholamandalam Coastline along the Bay of Bengal through the Utkala Kalinga region and end at the near the Sunderbans in Coastal East India.

2.1.7 North East

North east India is the eastern-most region of India connected to East India via a narrow corridor squeezed between Nepal and Bangladesh. It comprises the contiguous Seven Sister States—Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura—and the Himalayan state of Sikkim.

2.1.8 Island

It comprises of mainly Lakshadweep and Andaman and Nicobar islands. Lakshadweep formerly known as the Laccadive, Minicoy, and Amindivi Islands is a group of islands in the Laccadive Sea, 200 to 440 kilometres off the South Western

coast of India. The archipelago is a Union Territory and is governed by the Union Government of India.

The Andaman and Nicobar Islands are a group of islands at the juncture of the Bay of Bengal and Andaman Sea, and are a Union Territory of India. The territory is 150 km north of Aceh in Indonesia and separated from Thailand and Burma by the Andaman Sea. It comprises two island groups, the Andaman Islands and the Nicobar Islands, separated by the 10° N parallel, with the Andamans to the north of this latitude, and the Nicobars to the south. The Andaman Sea lies to the east and the Bay of Bengal to the west.

2.2 DISTRIBUTION OF SAMBAR, STUDY AREA AND SAMPLE COLLECTION

Sambar is distributed in six out of ten biogeographic zones including Himalaya, Semi-Arid, Western Ghats, Deccan Peninsula, Gangetic Plains and North-East. We have included the biological samples from three biogeographic zones in this study, which includes Western Ghats, Deccan Peninsula and Gangetic Plains. We collected biological samples from Madhya Pradesh, Goa, Kerala, Karnataka and Andhra Pradesh States (Table 2.2). Biological samples provided for wildlife forensics from various states were also used in this study (Table 2.2 and Figure 2.2). Tissue and faecal samples were collected in 70% ethyl alcohol and stored at 4°C till further analysis. Approximately, 4-6 inch basal portion of antler were collected from the field and stored at room temperature tills further analysis.

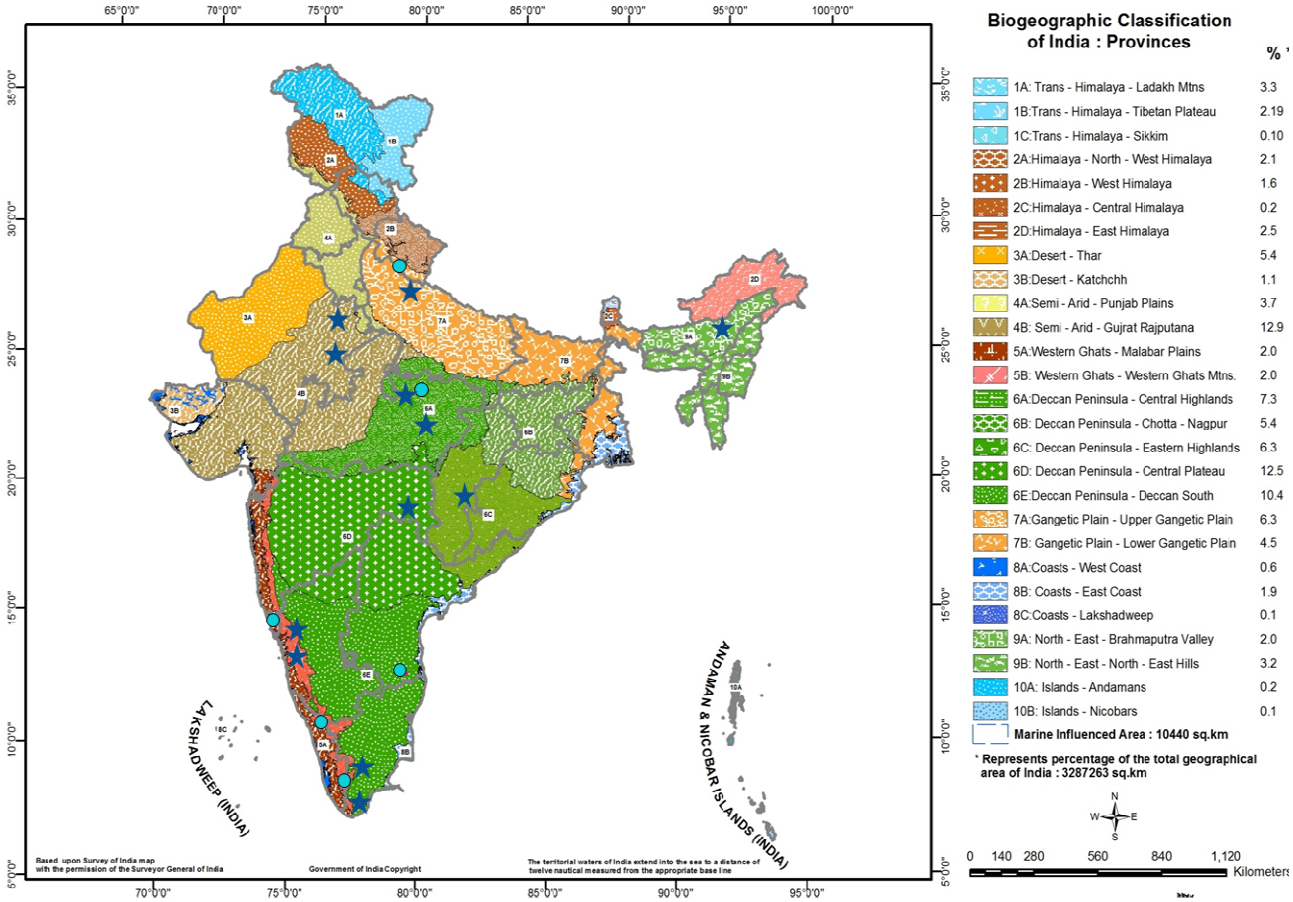


Figure 2.2: Geographical location of examined biological samples. Circles and stars represent location of non-invasive and forensic samples, respectively.

Table 2.2: Details of samples collected from distribution range from five bio-geographic zones of India

S.N.	Bio-geographic Zone	State	Protected Areas	No.
1	Deccan Peninsula	M.P. (Central highland), A.P. (Eastern highland)	Panna Tiger Reserve (TR), Shri Venkateshwara National Park (NP)	**45**
2	Western Ghats (Central and South)	Goa, Kerala	Bondla Forest Division, Periyar, Wayanad NP	**25**
3	Gangetic Plains (Upper Gangetic Plains)	Uttarakhand	Lansdowne Forest Division. and Corbet TR	**10**
4	Semi-Arid,	Rajasthan	Ranthambore & Sariska T.R's.	**4**
5	North-East	Assam	Kaziranga NP and Guwahati Zoo	**5**

2.3 METHODS

2.3.1 Assessment of various DNA extraction protocols

DNA extraction is the first and important step in the field of molecular genetics. More often, DNA extraction with higher PCR success rate is a challenging task from intricate biological samples including bone, antlers and fecal matters. The DNA extracted from these samples contains inhibitors which affect the PCR success (Rohland and Hofreiter, 2007a; 2007b). The bewildering situations for selecting the better DNA extraction protocol with higher PCR success rate is one of the frequent situation in conservation genetics. In the field of molecular genetics, the quantity, quality and the success rate of PCR amplification of extracted DNA has great importance (Gupta et al. 2011).

In conservation genetics, the uses of the non-invasive sample have the enormous scope and maximize the use of field effort in laboratory by increasing the sample size. Extraction of good quality of DNA from non-invasive samples including antler (natural shed), bone (from predator kills and natural death), and fecal material are a crucial task for the researchers to maximize the efficient use of field exercise. The silica based extraction DNA binding technique with chaotropic salts Guanidine thiocyanate (Gu-SCN) based lysis buffer is used for the extraction of DNA from hard

tissue like bone, teeth (Rohland and Hofreiter 2007a) and feces (Wehausen et al. 2004). The eluted DNA contains minimum inhibitory effect but the major drawback related to this protocol is the use of Gu-SCN, which is an expensive and hazardous reagent. This study highlights the more efficient use of less toxic and low-cost Guanidine hydrochloride (Gu-HCl).

Along with the use of chaotropic salts in commercial available Qiagen QIAamp DNA Stool Mini Kit/tissue kit (Qiagen Inc.), other effective DNA extraction methods including traditional phenol-chloroform (PC) protocol (Sambook et al. 1989) also provide good result. These protocols have been used on specific materials, and selection of the uniformly better amongst available protocols leads to great confusion. In this study, we have selected three different non-invasive biological samples and three different DNA extraction protocols to compare the quality of DNA. These three protocols were phenol-chloroform (PC), commercial kit and Guanidine hydrochloride (Gu-HCl) based in-house method. The extracted DNA samples were subjected to PCR amplification of 470 base pair long mtDNA fragment to demonstrate the effectiveness of various tested protocol (Gupta et al. 2013).

2.3.2 Processing of bone, antlers and fecal pallets

The antler and tissue samples of sambar, *Rusa unicolor* were used in this study. The fecal pallets of sambar were collected during this study from the field in either 70% ethanol or in silica gel. Starting materials with composition of calcium including bone and antlers requires decalcification before starting of DNA extraction by using 0.5 M EDTA. Bone and antlers were pulverized into small granule or pieces by using metal's motor and pestle. Decalcification of pulverized biological samples was done with 0.5 M EDTA, which requires 3-5 days incubation of sample with continuous rotation in a vertical rotor at room temperature replacement of fresh EDTA solution at regular interval of 8-10 hrs. Approximately 0.5 gm of the surface of fecal pallet was taken without any modification for each extraction.

2.3.3 DNA extraction

The first method used in this study was commercial available QIAamp DNA Stool Mini Kit/tissue kit (Qiagen, Germany) as described in product manual. The second method used was PC based DNA extraction protocol (Sambrook et al. 1989).

The third method used was Gu-HCl and silica binding DNA extraction protocol. 0.5 gm of above sample was mixed with 500 µl of lysis buffer (6 M Gu-HCl) and 20 µl of Proteinase-K (from 20mg/ml stock) in a 2 ml centrifuge tube. The tube was incubated at 56°C in hybridization over with continuous rotation for 24-48 hrs (till complete lysis). The tube was centrifuge at 13000 rpm for 3 min and clear solution was transferred to a fresh tube. 30 µl of glass milk (suspension of silica, SiO_2 powder in same volume of distilled water) was added in a fresh centrifuge tube and rotated in vertical rotor for 15 min for maximum binding of the DNA with glass particle. The tube was centrifuged at 14000 rpm for 2 min to get the glass pallet and the aqueous phase was discarded. 500 µl of wash buffer (20 mM Tris-HCL, pH 7.8, 1 mM EDTA, 50 mM NaCl, 50% ethanol) was added to the glass pallet and the same was mixed gently by pipette in and out for several times.

Subsequently, the tube was again centrifuged at 14000 rpm for 2 min to get the glass pallet and the aqueous phase was discarded. The washing step was repeated twice. The pallet was dried in heating block at 60 °C for 10 min. The glass pallet was suspended in 80 µl of TE buffer (10 mM Triss pH 7.8 and 1mM EDTA), the pellet was mixed gently by pipette in and put for several times. The tube was centrifuged at 14000 rpm for 5 min and clear aqueous DNA solution was transferred to a fresh tube (Gupta et al. 2013). Since the meat samples are a soft target for DNA extraction, hence; these were processed for PC based DNA extraction protocol (Sambrook et al. 1989) with minor modifications (Gupta et al. 2005) as and when needed.

2.3.4 PCR amplification

The DNA extracted from above three methods were used for PCR amplification of 472 bp long mtDNA Cyt *b* gene fragments and 580-642 bp long control (D-loop) region using following primer pairs:

2.3.4.1 Universal primers for mtDNA Cyt *b* gene

mcb398 ‘‘TACCATGAGGACAAATATCATTCTG’’ and
mcb869 ‘‘CCTCCTAGTTTGTTAGGGATTGATCG’’ (Verma and Singh 2003).

2.3.4.2 Deer specific primer for mtDNA control (D-loop) region

Cerv.tPro “CCACYATCAACACCCAAAGC” and
CervCRH “GCCCTGAARAAAGAACCAGATG” (Balakrishnan *et al*. 2003)

Several primers for PCR amplification of whole mtDNA cyt *b* gene were also designed and used in this study as given in table 2.3 (Gupta *et al*. 2014).

Table 2.3: Primers for amplification and sequencing of complete cyt *b* gene of ungulate species

S.No.	Name	Primer (Oligo) Sequence (5′ → 3′)
1	**MCB14081-105(F1)**	CATTATTCTCACATGGAATCTAACC
2	**MCB14560-584(F2)**	GAGGACAAATATCATTCTGAGGAGC
3	**MCB14584-560(R1)**	GCTCCTCAGAATGATATTTGTCCTC
4	**MCB14981-934(F3)**	ACCCAGACAACTACACCCCAGCAA
5	**MCB15934-981(R2)**	TTACTGGGGTGTAGTTGTCTGGGT
6	**MCB15347-323(R3)**	CTCCTTTTCTGGTTTACAAGACCAG

PCR reactions were carried out in 20µl reaction volume by using 1× PCR buffer (10 mM Tris-Hcl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 1× BSA, 100µM dNTPs, 4 pmol of each primers, 0.5 U AmpliTaq Gold DNA polymerase enzyme (Invitrogen Inc.) and 1µl (~10-40 ng) template DNA. The PCR conditions were: initial denaturation at 95°C for 10 min, followed by 35 cycles each of denaturation at 95 °C for 45 s, annealing at 55 °C for 40s, and extension at 72 °C for 1:30 min. The final extension was at 72 °C for 10 min. The efficiency and reliability of PCR reactions is monitored by using control reactions. Amplified fragments are confirmed by the electrophoresis on 2% agarose gel stained with ethidium bromide (0.5mg/ml) and visualized under U.V. transilluminator.

2.3.5 DNA sequencing

All positive PCR products were treated with exonuclease-I and shrimp alkaline phosphatase for 15 min. each at 37 °C and 80 °C, respectively, prior to sequencing PCR. After the BigDye terminator cycle sequencing reaction (BigDye RR 3.1 kit, ABI-Invotrogen).

2.3.5.1 Cycle sequencing PCR

For cycle sequencing PCR following component were added in reaction mixture.

COMPONENTS	PER REACTION
Big dye v 3.1 RR mix	0.25µl
5X sequencing buffer	2 µl
Primer	0.2µl
PCR product	1.0µl
Sterile water	6.55µl
Total Volume	**10.00 µl**

After preparing master mix using above component a PCR reaction was performed using following cyclic condition.

PCR condition for cycle sequencing-

Denaturation	96 ℃ for 0.10 seconds	30 cycles
Annealing	50℃ for 0.05 seconds	
Extension	60 ℃ for 4 minutes	

Stored at 4 ℃ or -20 ℃

2.3.5.2 Purification of sequencing reaction (clean up)

After completion of cycle sequencing PCR, the products were prepared for sequencing by alcohol precipitation method. This step was essential to clean the PCR product and remove the unused primer, DNA polymerase enzyme, dNTP's and labelled ddNTP's.

1. 40 µl of 24:1 alcohol: sodium acetate was added in each wells of PCR plate.
2. It was incubated at room temperature for 10 minutes
3. It was centrifuged at 4000 rpm for 20 minutes
4. The liquid was decanted carefully
5. 100 µl 70% ethanol was added in each wells
6. It was again centrifuged at 4000 rpm for 20 minutes
7. The liquid was again decanted carefully
8. It was dried for 5 minutes
9. 10 µl Hidi (Formamide) was added in each well.
10. Samples were mixed gently in HiDi and denatured at 95℃ for 3 minutes

Finally it was loaded in Applied Biosystem 3130 Genetic Analyzer for sequencing.

2.3.6 Analysis of sequencing data

Different softwares were used in sequence data analysis. These softwares are ClustalW (Thompson *et al*., 1994) provided in BioEdit package (Hall, 1999), DnaSP (Librado and Rozas, 2009), MEGA 5 (Tamura *et al.*2011), and the BEAST 1.7 (Drummond *et al.*2013). Details of the methodology and parameter used in these softwares are described in respective sections.

2.3.7 Genotyping

Twenty three STR loci were selected for assessment of genetic variation within various sambar populations (Table 2.4). All the forward (F) primers were labelled with fluorescent dye as given in table 2.4. After successful PCR amplification in multiplex reaction, the amplicons were used for genotyping by ABI 3130 Genetic Analyzer as using following protocol.

For genotyping 1 µl PCR product was mixed with 9 µl Hi-Di Formamide (*Life Technologies*) and 0.2 µl size standard Liz 500 (-250). This genotyping mix was then loaded in ABI 3130 DNA analyzer for obtaining the allele sizing data.

2.3.8 Statistical analysis of genotyping data

The data generated for 23 loci was first checked for any associated error and then subjected to further analysis. The heterozygosity measures were calculated using the appropriate software.

2.3.8.1 Genotyping errors

Following errors can commonly occur during amplification of microsatellite loci by polymerase chain reaction (PCR):

- Null alleles – one or more alleles fail to amplify during PCR.
- Stuttering – slight changes occur in the allele sizes during PCR.
- Large allele dropout – large alleles do not amplify as efficiently as small alleles.

Table 2.4: List of microsatellite markers used in this study

S. No.	STR Marker	Label	Allele Size	Reference
1	BM4208	NED	142-178	Bishop *et al.*1994
2	INRA011	VIC	148-277	Vaiman *et al.*1992
3	Ca18	6-FAM	128-212	Gaur *et al.*2003
4	RT1	VIC	234-256	Poetsch *et al.*2001
5	Cervid1	6-FAM	154-193	DeWoody *et al.*1995
6	ILSTS005	VIC	181-185	Brezinsky *et al.*1993
7	NVHRT48	VIC	105-115	Poetsch *et al.*2001
8	BM4107	NED	152-192	Bishop *et al.*1994
9	RT6	PET	93-109	Poetsch *et al.*2001
10	T108	6-FAM	128-178	Jones *et al.*2002
11	Ca42	PET	116-188	Gaur *et al.*2003
12	OarFCB193	VIC	96-124	Buchanan & Crawford 1993
13	D	NED	150-190	Jones *et al.*2000
14	T156	6-FAM	131-227	Jones *et al.*2002
15	T193	PET	172-238	Jones *et al.*2002
16	INRABERN185	VIC	244-250	Kappes *et al.* 1997
17	NVHRT16	NED	152-192	Poetsch *et al.*2001
18	BM6506	PET	184-200	Bishop *et al.*1994
19	T507	PET	143-187	Jones *et al.*2002
20	RT27	6-FAM	135-155	Poetsch *et al.*2001
21	T123	NED	146-178	Jones *et al.*2002
22	CelJP27	VIC	176-196	Marshall *et al.*1998
23	CSSM1 (AF232760)	PET	152-170	Moore *et al.*1992

The Hardy-Weinberg theory of equilibrium is used to calculate expected allele frequencies and the frequency of any null alleles detected.

2.3.8.2 Allele number

An allele is an alternative form of a gene that is located at a specific position on a chromosome. Allele number is the total number of alleles for a given marker/locus in a population, which is counted with non zero frequency. The allele number for each locus can be determined from genotype data using different software.

2.3.8.3 Allele frequency

Allele frequency is a measure of the relative frequency of an allele on a genetic locus in a population. Usually it is expressed as a proportion or a percentage. In population genetics, allele frequencies show the genetic diversity of a species population or equivalently the richness of its gene pool. The frequencies of all the alleles of a given gene often are graphed together as an allele frequency distribution histogram. The allele frequency is affected over time by forces such as genetic drift, mutation and migration. The frequency of alleles in a population can be used to predict the frequencies of the corresponding genotypes, in case of presence two alleles.

$p + q = 1$

$p^2 + 2pq + q^2 = 1$

Where 'p' is the frequency of one allele and q is the frequency of the alternative allele, which necessarily sum to unity. Then, p^2 is the fraction of the population homozygous for the first allele, 2pq is the fraction of heterozygotes, and q^2 is the fraction homozygous for the alternative allele. If the first allele is dominant to the second, then the fraction of the population that will show the dominant phenotype is p2 + 2pq, and the fraction with the recessive phenotype is q^2.

With three alleles:

$p + q + r = 1$

$p^2 + 2pq + 2pr + q^2 + 2qr + r^2 = 1$

In the case of multiple alleles at a diploid locus, the number of possible genotypes (G) with a number of alleles (a) is given by the expression:

$G = a\,(a+1)/2$

2.3.8.4 Effective Allele Number (Kimura and Crow, 1964)

The effective number of alleles (Ne) is the reciprocal of the sum of the square of allele frequencies.

$Ne = 1/\sum Pi^2$

Where Pi is the frequency of the i^{th} allele

2.3.8.5 Gene diversity (Heterozygosity)

Heterozygosity is the state of possessing different alleles at given locus in regard to given character. It is a measure of heterozygotes or genetic variation in a population. The population heterozygosity at a locus is given by the formula.

$H = 1 - \Sigma\, Pi^2$

Where, **Σ** a stand for summation over all alleles (Nei, 1978) and Pi is the frequency of the ith allele at a locus in a population. The average heterozygosity per locus (H) is defined as the mean of H over all structural loci in the genome.

2.3.8.4 Polymorphism Information Content (PIC)

Polymorphic Information Content (PIC), a measure of marker's informativeness was calculated according to Botstein *et al.,* (1980) using the given formula.

$$PIC = 1 - \sum_{i=1}^{j} Pi^{2} - 2\sum_{i=j+1}^{j} \sum_{J=1}^{i-j} Pi^{2}\, Pj^{2}$$

Where k is the number of alleles and χi, χj are the frequencies of the i^{th} and j^{th} alleles, respectively.

2.3.8.5 Hardy-Weinberg Equilibrium (HWE)

The Hardy–Weinberg principle states that both allele and genotype frequencies in a population remain constant—that is, they are in equilibrium—from generation to generation unless specific disturbing influences are introduced like non random

mating, mutations, selection and limited population size. The equilibrium is represented by the following equation:

$p^2 + 2pq + q^2 = 1$

and

$p + q = 1$

Where, p and q are the dominant and recessive allele frequencies of a given gene. Deviations from expected values may be due to the following reasons. If an excess of heterozygote is observed this may indicate the presence of over dominant selection or the occurrence of out-breeding. If an excess of homozygotes is detected it may be due to the presence of null alleles, inbreeding in the population, selection at the locus or Wahlund's effect i.e. reduction in observed heterozygosity when compared to expected.

2.3.8.6 F-statistics (Weir and Cockerham, 1990) and gene flow (Slatkin and Barton, 1989)

F statistics provide a tool for analysis of variance among populations. In population genetic studies, the use of F statistics enables the loss of variance to be partitioned into its constituent components allowing conclusion about the genetic structure of the population and relative levels of gene flow and inbreeding. A within subpopulation F-statistics can be determined from the ratio of observed to expected heterozygosity.

$F_{IS} = H_S - H_I / H_S$

Where H_S is the average heterozygosity estimated from each subpopulation.

$$H_S = 1 - \sum_{i=1}^{k} Pi^2$$

And H_I is the average observed heterozygosity

$$H_I = 1 - \sum_{i=1}^{k} H_I/k$$

For k subpopulations.

Among subpopulations F-statistics can be estimated from

$F_{ST} = H_T - H_S / H_T$

Where

$$H_T = 1 - \sum_{i=1}^{k} Pi^2$$

And Pi is the frequency of the i^{th} allele averaged over all subpopulations.

The measure of correlation of alleles for the entire population is thus a combination of both the within and among subpopulation effects, and can be estimated from

$F_{IT} = H_T - H_I / H_T$

The three values are related by the formula

$(1-F_{IT}) = (1- F_{IS}) (1- F_{ST})$

F_{IS} is concerned with inbreeding in individual (1) relative to the subpopulation (S) to which they belong, F_{ST} is concerned with inbreeding in subpopulation(s) relative to the total population (T) of which they are a part and F_{IT} is concerned with inbreeding in individuals (1) relative to the total population (T). The F- statistics can be estimated from observed frequencies. F_{ST} is related to migration by:

$F_{ST} = 1/4N_e\, m+1$

Where, N_e is the effective population size and 'm' is the effective proportion of immigrants. $N_e m$ gives a measure of gene flow across populations.

2.3.8.7 Population assignment

Genotype assignment for all pair of populations was performed by computing the log likelihood of the genotype of each individual in every sample, as if it was drawn from a population sample having allele frequencies equal to those estimated for each sample (Paetkau et al., 1997; Waser and Strobeck, 1998). Multi-locus genotype likelihoods are computed as the product likelihood of each locus assuming that the loci are independent. The assignment of individuals to the correct population and to other populations was also estimated.

2.3.8.8 Software used for data analysis

There were a wide range of statistical programs available for the analysis of molecular genetic data. Choosing the most appropriate program will depend on the

nature of specific data set and the final goal of the study. Following programs were used for data analysis.

i. **CONVERT** version 1.31 (Glaubitz, 2004) was used for inter conversion of different file format required for different genetic analysis software.

ii. **CERVUS** version 3.0 (Kalinowski *et al.* 2007) was used for predicting null allele frequencies and polymorphic information content.

iii. **GENEPOP** version 4.0 (Raymond and Rousset, 1995) was used for Hardy Weinberg Equilibrium estimation.

iv. **STRUCTURE** version 2.3.1 (Pritchard *et al.* 2000) was used for inferring Population assignment and individual ancestry using a clustering procedure.

CHAPTER 3
INTRA-SPECIES VARIATIONS AMONG SAMBAR POPULATIONS

SUMMARY

Despite its large distribution and significant morphological variation little study has been done to examine the genetic variation among sambar (*R. unicolor*) populations in India. This chapter describes genetic variation among select sambar populations using Mitochondrial DNA (mtDNA) control region. Sequence variation in a partial fragment of mtDNA control region was examined from the biological samples collected from Southern India, Central highlands and Semi-Arid regions. This study revealed that the south Indian population has a 40 bp insertion in control region. The populations of sambar found in few southern Indian sites, Central highland and Gangetic plains lack this feature. The molecular feature of 40 bp insertion indicates that these select sambar populations in the Western Ghats could probably be an ecologically significant unit (ESU).

3.1 BACKGROUND

Out of seven recognised subspecies of sambar, *R. u. unicolor* occurs in India (Leslie, 2011). Due to large distribution range in India, it would have praiseworthy to explore the level of genetic variation across different populations. Very limited genetic information was available for this species from its range. The objective of this chapter was to examine genetic variation amongst sambar populations in India.

Mitochondrial DNA (mtDNA) control region have been extensively used in population genetic study of several wild species e.g. tiger, *Panthera tigris* (Luo *et al.*, 2004); wild pigs, *Sus scrofa* (Larson *et al.*, 2005); Cervidae or deer (Douzery and Randi, 1997; Randi *et al*. 2001). Nucleotide sequence variation and variable copy number of a CR-I repeat, have been used to infer phylogenetic relationships among Cervides (Randi *et al*. 2001). Control region based phylogenetic framework depicted novel relationships and suggested that the formal taxonomy of the cervides and indicated that the genus *Cervus* needs modification (Randi *et al*. 2001). During this study, genetic variation within select Sambar population was examined using

sequence analysis of partial fragment of mtDNA control region (Balakrishnan *et al.*, 2003). The genetic variation of the sambar of Southern India was compared with that of Deccan Peninsula (Central highlands), Gangetic plains and Semi-Arid populations.

3.2 MATERIALS AND METHODS

3.2.1 Starting material and DNA extraction

A total of 59 biological samples including 53 tissue and six fecal samples were used in this study. 22 samples were from south and 37 from central highland and north India. In India wildlife enforcement agencies often confiscate the biological samples of wild species from poachers for appropriate punishment by court of law. These samples need to be validated by laboratories for species identification before the suitable punishment. 35 validated biological forensic samples of sambar were also used for this study. Biological samples were also collected from national park and captive populations of sambar after obtaining permission from the competent authorities. Six tissue and one antler samples were collected from the remnant of predator kills from Panna Tiger Reserve, Madhya Pradesh. Four and seven hair samples were collected from confined populations from deer parks of Goa and Andhra Pradesh States, respectively. Six fecal samples were collected from the forest of Kerala State. Whole genomic DNA was extracted using modified Phenol-Chloroform (Sambrook *et al*., 1989) and indigenous protocol (Gupta *et al*., 2013).

3.2.2 PCR Amplification

The DNA extracted was used for PCR amplification of approximately 650 bp long mtDNA control region. The primer pair used for PCR amplification is specific for Cervidae (deer) family Cerv.tPro "CCACYATCAACACCCAAAGC" and CervCRH "GCCCTGAARAAAGAACCAGATG" (Balakrishnan *et al*., 2003). PCR reactions were carried out in 20µl reaction volume by using 1× PCR buffer (10 mM Tris-Hcl, pH 8.3, and 50 mM KCl), 1.5 mM $MgCl_2$, 1× BSA, 100µM dNTPs, 4 pmol of each primers, 0.5 U AmpliTaq Gold DNA polymerase enzyme (Invitrogen Inc.) and 1µl (~30 ng) template DNA. The PCR conditions were: initial denaturation at 95°C for 10 min, followed by 35 cycles of denaturation at 95°C for 45s, annealing at 54°C for 40s, and extension at 72°C for 75s. The final extension was at 72°C for 10 min. The efficiency and reliability of PCR reactions is monitored by using control reactions. Amplified fragments are confirmed by the electrophoresis on 2.20% agarose

gel stained with ethidium bromide (0.5mg/ml) and visualized under U.V. transilluminator (Figure 3.1).

3.2.3 DNA sequencing and analysis

Before sequencing all PCR products were treated with exonuclease-I and shrimp alkaline phosphatase for 15 minute each at 37°C and 80°C. After the BigDye terminator cycle sequencing reaction and purification the samples were sequenced by 3130 ABI automated DNA genetic analyzer (Applied Biosystem) from both the directions. The sequences generated were aligned by eye using ClustalW (Thompson *et al*., 1994) provided in BioEdit package (Hall, 1999). The primer pair used in this study amplifies approximately 40-50 bases of Proline tRNA at the beginning of PCR reaction. Therefore, initial sequence of Proline tRNA was removed from the aligned sequences and hence, start sequences used in this study are first nucleotide of control region. Aligned sequences of control region were used for haplotype analysis using DnaSP (Librado and Rozas, 2009) (Table 3.1). To understand the difference within the sequence groups (haplotypes), a percentage similarity matrix and pair wise distance matrix were generated using ClustalW (Thompson *et al*., 1994) and MEGA 5 (Tamura et al 2011) programs, respectively (Table 3.2).

3.3 RESULTS

Approximately 40 bp longer PCR amplifications were obtained with the DNA extracted from the biological sample of Western Ghats (Fig. 3.1). Biological samples of Deccan Peninsula and Gangetic plains yielded expected size (~540 bp) of amplification. This indicated that the Western Ghats population has unique feature in mtDNA control region. Using the primer pair 543-623 bp sequences of control region were obtained, which were aligned for further analysis. Upon alignment, all the sequences generated from South India samples exhibited 40 bp insertion after nucleotide (nt) position 233 (Table 3.1).

This unique insertion in Western Ghats population (N = 22) indicated that these population have significant genetic variation. Overall mtDNA control region was highly variable among sambar populations and exhibited 26 haplotypes (Table 3.1). All the haplotype (RUC1-26) sequences were submitted to GenBank, and their accession number were KF133981-99 and KF648589-95 (Table 3.1). Haplotypes

RUC1-5, 8-15, 20 and 21 are characteristics of Western Ghats population. Most frequent haplotype was RUC17, which was observed in 24 individuals of Deccan Peninsula, Gangetic Plain and Semi-Arid populations. Haplotypes RUC2 and 18 was observed in four and RUC19 in three individuals of Deccan Peninsula (Table 3.1). Pair wise distance matrix indicated a distance range of 0.002-0.113, whereas percentage similarity matrix showed a very wide range (91.60-99.83) (Table 3.2). Insertion positive and negative sites are illustrated on the map if India as shown in Figure 3.2.

Deletion observed in couple of the individuals sample collected from Kalakkad Mundanthurai Tiger Reserve (KMTR), Tamil Nadu (Western Ghats) indicating that these area could be probable site for existence of insertion negative population also which probably could have migrated from Deccan Peninsula.

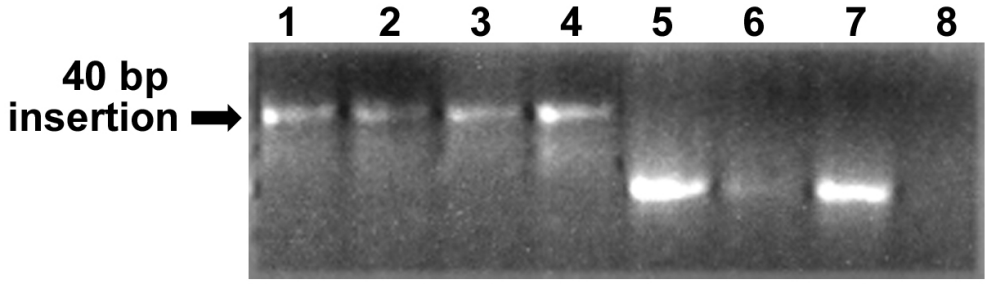


Figure 3.1: Gel image showing 40bp longer PCR amplification from the samples of South India (lane 1-4) and shorter amplification from the samples of Central highland, Gangetic plain and south India (lane 5-7)

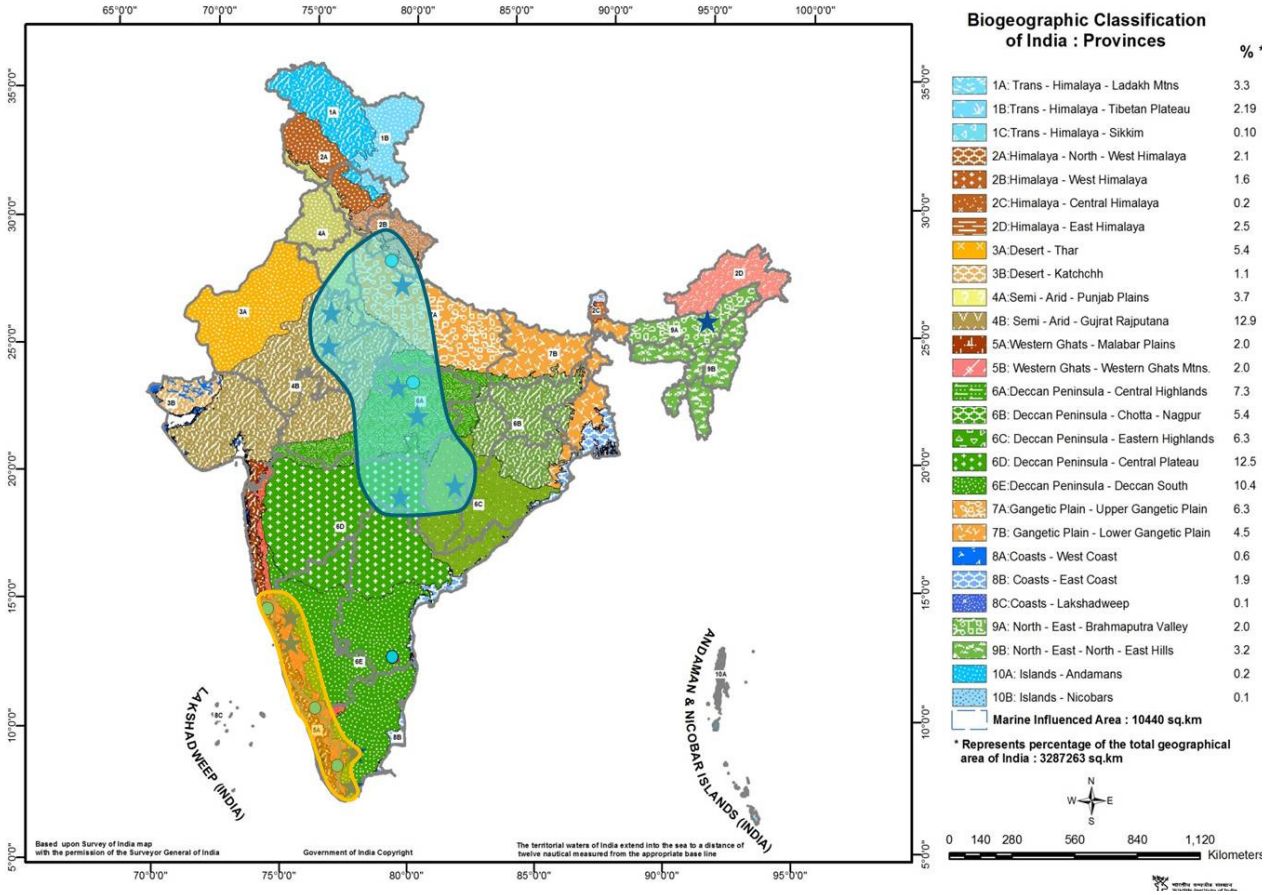


Figure 3.2: Locations of INDEL sites on map of India. The areas represented by yellow and blue shed indicate the locations of 40 bp insertion and deletion, respectively in mtDNA control region.

Table 3.1: Table showing different haplotype (RUC1-26) with number of repeat in square bracket [ ]. 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. Gray shades represent the 40 bp gap.

Haplotype	No. of Sample	GenBank Acc. No	3 1	6 5	5 8	1 1 9	1 2 0	1 2 1	1 2 2	1 2 8	1 3 8	1 4 0	1 4 1	1 4 2	1 5 1	1 5 2	1 5 3	1 5 5	1 5 6	1 5 7	1 5 8	1 5 9	1 6 1	1 6 4	1 7 0	1 7 4	1 7 5	1 7 8	1 7 9
RUC1	[1]	KF133981	G	C	A	T	T	A	G	T	G	-	-	-	G	T	C	G	T	C	T	A	A	A	A	A	A	T	A
RUC2	[4]	KF133982	.	.	-	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RUC3	[1]	KF133983	.	.	-	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RUC4	[1]	KF133984	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.
RUC5	[1]	KF133985	.	.	.	.	.	.	.	.	.	-	-	-	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
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
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	.
RUC10	[1]	KF133990	A	T	.	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC11	[1]	KF133991	A	T	.	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC12	[1]	KF133992	A	T	.	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC13	[1]	KF133993	A	T	-	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC14	[1]	KF133994	A	T	-	C	-	-	A	A	A	G	T	T	A	.	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC15	[1]	KF133995	A	T	-	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC21	[1]	KF648590	A	T	-	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	G	.	G	.	.	.
RUC16	[1]	KF133996	A	T	.	C	T	A	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	.	.	.	G	.	G
RUC17	[24]	KF133997	A	T	.	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G
RUC18	[4]	KF133998	A	T	-	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G
RUC19	[3]	KF133999	A	T	-	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G
RUC6	[2]	KF133986	A	T	-	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G
RUC7	[1]	KF133987	A	T	-	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	G	G	.	G	.	G
RUC22	[1]	KF648591	A	T	.	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	G	G	.	G	.	G
RUC23	[2]	KF648592	A	T	-	C	-	-	A	A	A	G	C	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC24	[1]	KF648593	A	T	-	C	-	-	A	A	A	G	T	T	A	C	G	A	C	T	C	.	.	.	.	G	.	.	.
RUC25	[1]	KF648594	A	T	-	C	-	-	A	A	A	G	T	T	.	C	G	A	C	T	C	G	G	G	G	.	G	.	G
RUC26	[1]	KF648595	A	T	.	C	-	-	A	A	A	G	C	T	.	C	G	A	C	T	C	.	G	.	.	.	G	.	G

Haplotype	No. of Sample	GenBank Acc. No.	18	18	19	19	19	19	19	20	20	20	21	22	22	22	22	22	22	22	22	23	23	23	23	24
			38	38	01	01	27	27	99	00	33	99	99	00	33	00	37	60	00	77	89	20	10	99	00	06
RUC1	[1]	KF133981	A	C	C	A	T	G	C	C	A	A	A	T	G	T	T	T	C	T	T	C	T	T	C	T	A
RUC2	[4]	KF133982	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	C	.	.	.	.
RUC3	[1]	KF133983	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	C	.
RUC4	[1]	KF133984	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.
RUC5	[1]	KF133985	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
RUC20	[2]	KF648589	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.
RUC8	[1]	KF133988	G	T	T	.	C	A	.	T	.	.	.	.	.	.	.	C	.	T	.	A	T	.	A	T	C	.
RUC9	[1]	KF133989	G	T	T	.	C	A	T	T	.	.	.	.	.	.	.	.	.	T	.	A	T	.	A	T	C	.
RUC10	[1]	KF133990	G	T	.	G	.	A	.	T	G	G	.	.	T	C	.	.	T	.	A	T	.	A	.	.	.
RUC11	[1]	KF133991	G	T	.	G	.	A	.	T	G	G	.	C	T	C	C	.	T	.	A	T	.	A	T	.	.
RUC12	[1]	KF133992	G	T	.	G	.	A	.	T	G	G	.	C	T	C	.	.	.	.	A	T	.	A	T	.	.
RUC13	[1]	KF133993	G	T	.	G	.	A	.	T	G	G	.	.	T	C	.	.	.	.	A	T	C	A	T	.	.
RUC14	[1]	KF133994	G	T	.	G	.	A	.	T	G	G	.	.	T	C	.	.	.	.	A	T	C	A	T	.	.
RUC15	[1]	KF133995	G	T	.	G	.	A	.	T	G	G	.	.	T	C	.	.	.	.	A	T	C	A	T	.	.
RUC21	[1]	KF648590	G	T	.	G	.	A	T	T	G	G	.	.	T	.	.	.	T	.	A	T	.	A	T	.	.
RUC16	[1]	KF133996	G	T	T	.	.	A	T	T	.	.	.	.	.	.	-	-	-	.	A	T	.	A	T	C	.
RUC17	[24]	KF133997	G	T	T	.	C	A	T	T	.	.	.	C	.	.	-	-	-	.	A	T	.	A	T	.	.
RUC18	[4]	KF133998	G	T	T	.	C	A	T	T	.	.	.	.	.	.	-	-	-	.	A	T	.	A	T	.	.
RUC19	[3]	KF133999	G	T	T	.	C	A	T	T	.	.	.	.	.	.	-	-	-	.	A	T	.	A	T	.	G
RUC6	[2]	KF133986	G	T	T	.	C	A	T	T	.	.	.	.	.	.	-	-	-	.	A	T	.	A	T	.	.
RUC7	[1]	KF133987	G	T	T	.	C	A	.	T	.	.	.	.	.	.	-	-	-	C	A	T	.	A	T	.	.
RUC22	[1]	KF648591	G	T	T	.	C	A	T	T	.	.	.	.	.	.	-	-	-	.	A	T	.	A	T	.	.
RUC23	[2]	KF648592	G	T	.	G	.	A	.	T	G	G	.	.	T	C	-	-	-	.	A	.	C	A	T	.	.
RUC24	[1]	KF648593	T	T	.	.	.	A	T	T	.	.	.	.	.	C	-	-	-	.	A	T	C	A	T	.	.
RUC25	[1]	KF648594	G	T	T	.	C	A	T	T	.	.	.	.	.	.	-	-	-	-	-	-	-	A	T	.	.
RUC26	[1]	KF648595	G	T	T	.	.	A	T	T	.	.	.	.	.	.	-	-	-	-	-	-	-	A	T	C	.

Haplotype	No. of Sample	GenBank Acc. No.	4 0 8	4 1 4	4 3 0	4 3 3	4 3 6	4 3 8	4 5 0	4 5 1	4 6 7	4 6 9	4 7 2	4 8 3	4 9 6	4 9 9	5 1 0	5 1 1	5 3 0	5 4 3	5 7 6	5 7 9	5 8 0	6 0 3
RUC1	[1]	KF133981	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	.	.	A	.	.	.	T	.	G	.	.	A	C	G	.	.	.	A	.	.	.	.
RUC3	[1]	KF133983	.	.	.	.	.	.	T	.	G	.	T	A	C	G	.	.	.	A	A	.	.	.
RUC4	[1]	KF133984	.	.	A	.	.	.	.	.	G	.	.	A	C	G	.	.	.	A	A	.	.	T
RUC5	[1]	KF133985	.	.	A	.	.	.	.	A	.	.	.	.	.	.	G	G	T	.	.	.	C	.
RUC20	[2]	KF648589	.	.	.	.	.	.	T	.	G	.	T	A	C	G	.	.	.	A	A	.	.	.
RUC8	[1]	KF133988	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC9	[1]	KF133989	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC10	[1]	KF133990	C	C	A	.	C	T	T	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC11	[1]	KF133991	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	T
RUC12	[1]	KF133992	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC13	[1]	KF133993	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC14	[1]	KF133994	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC15	[1]	KF133995	C	.	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC21	[1]	KF648590	C	C	A	C	C	.	.	.	.	T	T	.	.	.	G	.	.	.	.	.	C	.
RUC16	[1]	KF133996	C	.	A	.	.	.	.	A	.	T	T	A	.	.	.	.	.	.	A	C	.	.
RUC17	[24]	KF133997	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC18	[4]	KF133998	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	.	T	.	A	C	C	.
RUC19	[3]	KF133999	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC6	[2]	KF133986	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC7	[1]	KF133987	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC22	[1]	KF648591	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC23	[2]	KF648592	C	C	A	.	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC24	[1]	KF648593	C	C	A	C	C	T	.	.	.	T	T	.	.	.	.	G	.	.	.	.	C	.
RUC25	[1]	KF648594	C	C	A	C	C	T	.	A	.	.	.	.	.	.	G	G	T	.	A	C	C	.
RUC26	[1]	KF648595	C	.	A	.	.	.	.	A	.	T	T	A	.	.	.	.	.	.	A	C	.	.

Table 3.2: Table showing pair wise distance (above diagonal) and percent similarity (below diagonal) amongst control region sequences generated by deer specific primer (Balakrishnan et al., 2003).

	RUC1	RUC2	RUC3	RUC4	RUC5	RUC2	RUC8	RUC9	RUC1	RUC1	RUC1	RUC1	RUC1	RUC1	RUC2	RUC1	RUC1	RUC1	RUC1	RUC6	RUC7	RUC2	RUC2	RUC2	RUC2	RUC2
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	94.65	94.97	-	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	94.65	94.97	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	94.81	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	94.81	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	94.65	94.97	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	94.65	94.97	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	94.97	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	94.55	96.4	96.06	94.89	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	94.7	94.35	94.18	94.18	94.35	94.35	94.52	-	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	94.96	94.96	94.96	94.59	94.4	94.77	96.14	99.08	96.14	96.14	94.77	94.96	94.77	96.32	94.04	94.22	96.49	-

3.3.1 Photographic variations

During field work, photographic images of sambar were collected from the forest area of Western Ghats, Central highland and Gangetic plains for comparing photographic feature. Upon comparison of photographic feature, it was observed that, all Indian sambar populations have very similar body feature and antler growth. However, sambar populations found in selected area in southern India (in Western Ghats) are having active sore spot on their neck region (Plate 2). This patch becomes more prominent in rut and causing reddish oedema like appearance. It sometimes oozes a white liquid, and is apparently glandular in nature.

This reddish oedema like glandular feature, which is an important morphological variation in sambar, could be an ecological distinction of these populations with those which do not have this feature. This unique and interesting morphological feature of select southern India population was observed during this study. It was also found that all sambar do not have active sore spot, however; several sambar populations in South India have sore spot. Researchers have observed active sore spot in few North-east India populations; however, it has not been reported in literature. Previous mistaken report suggested that entire sambar populations have sore spot (Geist 1998). However, during this study, I observed that the sambar populations found in several places in south India have this feature (Plate 2 and 3). Extensive sampling from north-east India could not be performed due to lack of permission. Other sambar populations inhabiting in Central highland, Gangetic plains and Semi-Arid regions appeared to not possessing an active sore spot (Plate 4).

Plate 2: Image of sambar hind during breeding season in Western Ghats (Parambiculam Tiger Reserve, Kerala).

Plate 3: Image of sambar stag in forest area of Western Ghats (Bondla Division, Goa)

Plat 4: Image of sambar in Central highland (Panna Tiger Reserve, Madhya Pradesh)

3.4 DISCUSSION

MtDNA control region were used in phylogeny studies of wild species (Larson *et al*. 2005; Balakrishnan *et al*. 2003). This region has been implemented to confer that the critically endangered Eld's deer (*Rucervus eldi eldi*) is more closely related to *R. e. thamin* than to *R. e. siamensis* (Balakrishnan *et al*. 2003). In this study, control region was used to evaluate the genetic variations among sambar population in India. An interesting Insertion–deletion (INDEL) was observed among studied sambar population. INDEL markers were commonly used in the human genetics, which has numerous advantages for population and forensic studies, such as easy interpretation, small amplicons and easy genotyping (Thangaraj *et al*., 2006; da Costa Francez *et al*., 2012). The same primer pair which is used in this study had also been used in Eld's deer and 78 bp insertion was observed in selected population of Eld's deer (*Rucevus eldi*), which further exhibited ecological significance in understanding the difference in populations (Balakrishnan *et al*. 2003). This region has also been used in population genetics of several wild animals including the tiger, *Panthera tigris* (Luo *et al.* 2004, Sharma *et al.* 2009) and wild pig, *Sus scrofa* (Kim *et al.* 2002; Larson *et al.* 2005). Present chapter described the use of control region for differentiation of genetically distinct sambar population. INDEL of 40 bp long stretch was detected in the same region in the sambar populations of geographically distinct locations across India (Figure 3.2). This unique 40 bp insertion was detected in select population of Western Ghats, which was absent in the sambar populations of Gangetic Plains, Semi-Arid, Deccan Peninsula and few individuals of area located at the fringes of Western Ghats and Deccan Peninsula. This unique molecular feature in the sambar populations of Western Ghats differentiates them from remaining populations of India. It indicates that these populations could probably be significant for management interventions and further evolutionary study. This study also scientifically highlighting the existence of one ecologically important molecular signature (40 bp insertion at nt position 233) in select sambar population. This chapter widely covered the sambar populations from North to South India and highlighted a differentiable molecular feature of 40 bp insertion at nt position 233. Furthermore, the genetic variation level was estimated up to 8% (Table 3.2), which was significantly a higher distance for intra-species individuals.

As expected faunal radiation was favoured towards peninsular regions (Mani 1974), individuals from northern regions would have migrated towards the southern boundary of Deccan Peninsula. Therefore, deletion was observed in couple of the individuals sampled from Kalakkad Mundanthurai Tiger Reserve (KMTR), Tamil Nadu (Western Ghats), which is situated at the fringe of Western Ghats and Deccan Peninsula (Deccan south). It also indicated that this area could probable be one of the sites for existence of insertion negative population which might have migrated from Deccan Peninsula.

In assessment of population level genetic variations among wild species, control region has been proven (Luo *et al.* 2004; Kim *et al.* 2002; Larson *et al.* 2005). Sequences from complete mtDNA control regions were used to infer phylogenetic relationships in 25 Cervidae taxa (Randi *et al.* 2001). Based on control region, it was indicated that *Cervus elaphus* comprise two divergent clades those must be referred to as species *elaphus* (European elaphoid deer) and *canadensis* (Eurasian and North American wapitoid deer). Whereas, *Cervus nippon* was divided into Japanese and continental plus Taiwan sika. It was also observed that Père David's deer was nested within *Cervus*, and hence suggested that *Elaphurus* should be merged with *Cervus*. European and Persian fallow deer were genetically divergent and distinct species (Randi *et al.* 2001). Randi *et al.* (2001) also indicated that control region has significant implication in revising the taxonomy of cervides. The existence of 40 bp insertion in control region in the selected Western Ghats population could be similar an indication. More than one haplotypes were observed from one population during this study (Table 3.1). The coexistence of distinct haplotypes in a population could probably be because of reduction of female–mediated gene flow due to recent habitat fragmentation and subsequent loss (Wu *et al.* 2006). Same region was useful in genetic differentiation among Eld's deer subspecies, and assisted in identification of three distinct evolutionarily significant units (ESUs) for proper wildlife management (Balakrishnan *et al.* 2003). In this study, two genetically distinct populations of sambar with excess of haplotype variations (total 26 haplotype) were identified for proper wildlife management.

3.5 CONCLUSIONS

Control region among sambar population was highly polymorphic. Detection of 26 haplotype indicated that sambar populations are genetically diverse in its habitats. INDEL of 40 bp long stretch was a significant finding of this chapter. Presence of 40 bp insertion in select Western Ghats populations indicated that this population could probably be significant for management interventions and further evolutionary study. This insertion was apparently not detected in the sambar populations of Gangetic Plains, Semi-Arid, Deccan Peninsula and few individuals of area located at the fringes of Western Ghats and Deccan Peninsula. This unique molecular feature appeared to be a potential molecular signature for genetic separation the two distinct sambar populations.

CHAPTER 4
MOLECULAR SYSTEMATIC AND ORIGIN OF SAMBAR

SUMMARY

This chapter describes the phylogenetic status of sambar deer, *Rusa unicolor*. The sequences of mitochondrial DNA cytochrome *b* (cyt *b*) gene were generated from Sambar population across India. These sequences were aligned and compared with that of the available cyt *b* sequences of sambar and other related cervides from GenBank. Phylogenetic analysis indicated that there was a wide distribution of one common ancestor of sambar around 2.8 million year ago (mya). This ancestor was probably fragmented and restricted to Western Ghats in India and in South-east Asia close to Javan islands. It was appeared that the modern sambar (*Rusa unicolor*) evolved from the south-east Asian ancestor and subsequently radiated towards other part of south Asia. The sambar population reached to north-east India and further moves across the India and accumulated recent genetic changes. The Western Ghats sambar apparently did not undergo evolutionary differentiation and retained its habitat. Hence, it appeared India has two distinct lineage of sambar one similar to south-east Asian lineage and other is genetically distinct Western Ghats lineage. Moreover, the Western Ghats sambar exhibit > 4% variations in cyt *b* gene, which genetically differentiate them from remaining sambar subspecies. Hence, Western Ghats sambar population could be ancient and significant evolutionary unit (SEU).

4.1 BACKGROUND

The sambar can be distinguished from other species of cervides by their robust antlers with a long acutely angled brow tine, very deep lacrimal pits, reduced auditory bullae, and dark eumelanic pelage (Groves and Grubb 1987). Two recognized *Rusa spp.* i.e. *R. unicolor* and *R. timorensis* are morphologically identical. At first look, it is difficult to differentiate between sambar and Javan sambar (rusa), but closer observation indicates that the Rusa's body weight is lesser than sambar and their ears are more pointed, while sambar's ears are more rounded. Additionally, the third or top tine of the rusa forks to the front of the main beam, while the sambar's forks to the back (Summers, http://www.bowhuntingsafari.com/SPrusadeer.html).

Of the seven recognized subspecies, *R. u. unicolor* is reported from India and Sri Lanka (Leslie 2011), which has marked variation in their body size and antler size (Lydekker and Dollman, 1985). Thus, its large distribution range and occurrence of multiple subspecies within closely related landscapes needs verification in terms of its phylogeography at molecular levels.

In resolving the phylogenetic relationships between closely related taxa genes located on mitochondrial DNA (mtDNA) are very useful marker (Moritz *et al.* 1987). Its high mutation rate and short coalescent time make it more suitable than nuclear DNA for evolutionary studies (Zink and Barrowclough 2008). Molecular phylogeny based on use of multiple loci is preferred (Chatterjee *et al.* 2009; Rakotoarisoa *et al.* 2010), however; use of single gene would be more beneficial because it standardized the loci, and reduce cost, time and complexity of comparison (Tobe *et al.* 2010). Different rates of evolution at different regions of the mitochondrial genome facilitate the selection of suitable regions for the appropriate question of a study (Saccone *et al.* 1991). Few most important regions of mtDNA are the gene for cytochrome *b* (cyt *b*), Cytochrome Oxidase subunit-I (CO*I*), 12s rRNA and 16s rRNA, which are commonly used in phylogenic study (Miyamoto *et al.* 1990; Randi *et al.* 2004).

Recent study based on comparison of common molecular markers used in phylogenetic study shows that cyt *b* gene provides better resolution when species are separated on the basis of sequence data (Tobe *et al.* 2010). It is expected that two members of same species have close to 100% similarity at cyt *b* gene. Match can be less than 100% in case of either intraspecific variation or the compared sequence comes from closely related species (Tobe *et al.* 2010). It has been established that these regions can be amplified in various species, using a single pair of universal primers by PCR reaction (Kocher et al. 1989; Parson *et al.* 2000; Heish *et al.* 2001; Verma and Singh, 2003).

Morphological characteristics such as antler structure were used in classification of the cervidae taxa, however; uncertainty in taxonomic position has been addressed by the use of cyt *b* marker (Pitra *et al.* 2004). In this chapter, molecular insight in to phylogenic status of sambar across different biogeographic zones of India was described using sequence analysis of cyt *b* gene.

4.1.1 Fossil evidence for origin of sambar

During the upper Pliocene, rusine deer were found in Europe (Di Stefano and Petronio 2002; Lydekker 1898; van Bemmel 1974), and early lower Pleistocene forms with 3-tine antlers, such as the Philis deer (*Cervus philisi*), have purported affinities with living Rusa (Kurten 1968; Lister 1987). *R. unicolor* is among the most ancestral of living cervids, with characteristics little changed from the late Pliocene and paralleling other Chinese pliocervines (Petronio *et al.* 2007). It likely evolved in southern tropical areas (Flerov 1952), perhaps from the extinct Pleistocene forms such as *Epirusa hilzheimeri* (Di Stefano and Petronio 2002; Zhdanski 1925) or Eucladoceros (Geist 1998; Grubb 1990; Koizumi et al. 1993). Di Stefano and Petronio (2002) proposed that *Rusa elegans* branched in the mid-Villafranchian, 2.0–2.5 million years ago, giving rise to the extinct *R. hilzheimeri*, which gave rise to *R. unicolor* and the high-elevation specialist *Przewalskium albirostre* of the Tibetan Plateau (Leslie 2009; Schaller 1998).

Five rusine fossil species (*Rusa elegans, R. microta, R. stehlini, R. unicolor* and *R. yunnanensis*) of early Pleistocene origin have been found throughout China (Dong 1993). Fissile record apparently indicated that *R. timorensis* did not evolve until the late Pleistocene. A fossil Rusa of middle Pleistocene 0.8–1.7 mya was found in caves in northern Vietnam, which was comparatively larger than extant *R. unicolor* (Bacon et al. 2004). All these fossil records indicate that sambar evolved during Pleistocene and it is a recent descendent of cervidae family. In this chapter, cyt *b* gene sequences were used to examine the molecular phylogeny of sambar.

4.2 MATERIAL AND METHODS

4.2.1 Starting material and DNA extraction

86 tissue and 11 fecal samples of sambar used for this study. These samples were collected from different locations of India (Fig. 2.2). Of these, 21 samples were from Western Ghats; 76 from Deccan Peninsula, Semi-Arid, North East and Gangetic Plains region. The whole genomic DNA was isolated by indigenous protocols (Gupta *et al*. 2013) and DNA extraction kits (DNeasy Tissue "QIAGEN Kit"). The detail methods of extraction and amplification of DNA has been given in chapter 2 section 2.3.

4.2.2 PCR amplification

Complete cyt *b* gene from degraded (extracted from putrefied tissue) and high quality (extracted from fresh tissue) DNA was amplified using a series of primer set developed for ungulates (Table 2.3). In addition, a partial fragment of cyt *b* gene was amplified from highly degraded DNA extracted from fecal samples using universal primer (Verma and Singh, 2003). The amplification was carried out in 20 µl reaction volume containing 20-40 ng of the template DNA, 100 µM each of dNTPs, 4 pmol of each primer, 1.5 mM $MgCl_2$, 0.5 units of AmpliTaq Gold (Perkin-Elmer-Cetus, USA), and 1 X PCR buffer (10 mM Tris–HCl, pH 8.3, and 50 mMKCl). The PCR conditions were: an initial denaturation at 95°C for 10 min, followed by 35 cycles each of denaturation at 95°C for 45 s, annealing at 55°C for 1 min and extension at 72°C for 1 min. The final extension condition was 72°C for 10 min. The PCR products were electrophoresed on 2% agarose gel and visualises under Ultra-Violet (UV) light in presence of ethidium bromide.

4.2.3 Post PCR cleanup and DNA sequencing

The 1.5 µl of each PCR product was incubated with 0.5 µl ExoSAP-IT (USB, Cleveland, Ohio) at 37°C for 20 min. followed by inactivation of enzymes at 85°C for 15 min. The PCR products after *Exo-I* and *SAP* treatment were sequenced using ABI 3130 Genetic Analyzer on both strands. The sequences generated for both the strands were examined and resolved for any ambiguity by using Sequencher 4.1 programme.

4.2.4 Sequence analysis

DNA sequences obtained were aligned by ClustalW (Larkin *et al.* 2007) programme along with the DNA sequences of closely related deer which were either generated by us or downloaded from NCBI GenBank. The aligned sequences used for construction of maximum likelihood (ML) tree at bootstrapping for 1000 replications to get a precise tree by MEGA 5 programme (Tamura *et al.* 2011). Furthermore, a phylogenetic tree using Markov Chain Monte Carlo (MCMC) based Bayesian approach was constructed by BEAUti and the BEAST 1.7 (Drummond et al 2013). A matrices for percentage similarity and pairwise distance were generated using ClustalW (Larkin *et al.* 2007) and MEGA 5 programme (Tamura *et al.* 2011), respectively (Table 4.1).

4.3 RESULT

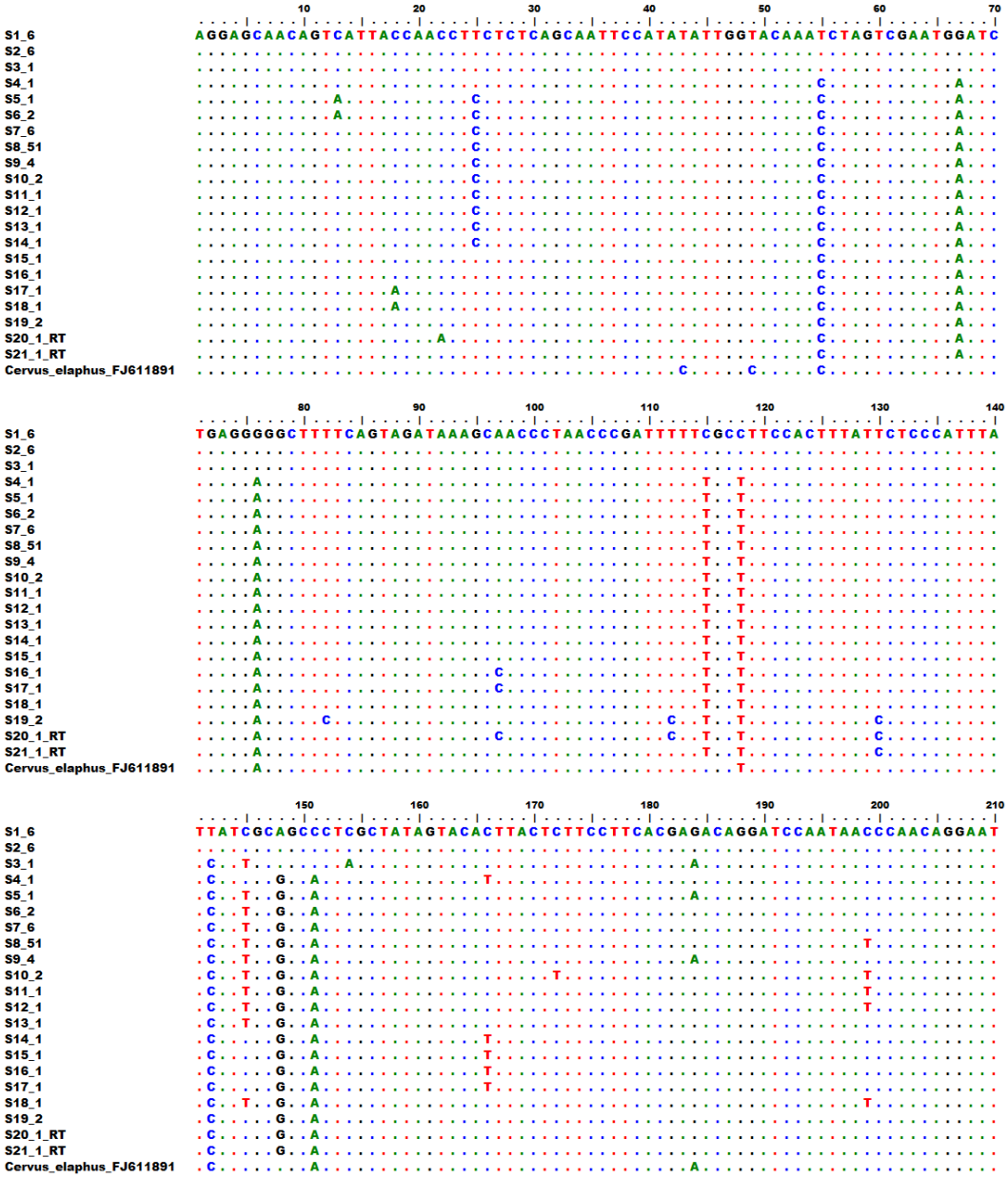
Based on similarity among 77 in-house and 11 GenBank sequences of *R. unicolor* and two sequences of Javan sambar (*R. timorensis*), 21 haplotype were identified using DnaSP program and named as S_{1-21} (Table 4.1). The haplotype S_{1-19} were of recognized *R. unicolor* and S_{20-21} were of *R. timorensis* (Fig. 4.1). While examining the sequence variation in a partial fragment of cyt *b* of all sambar samples, an interesting disparity in select sambar population was observed in Western Ghats.

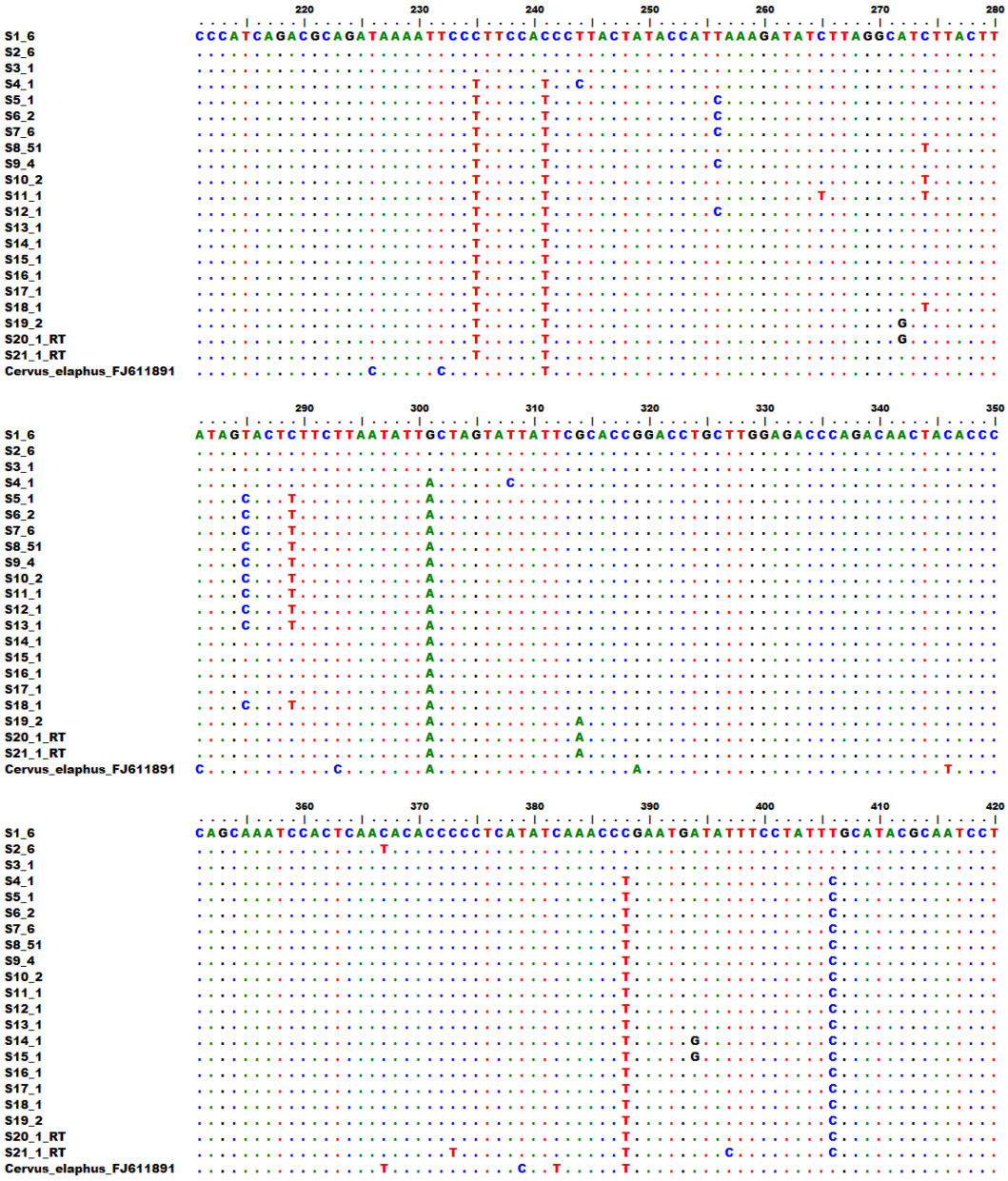
Three haplotypes labeled as S_{1-3} were exclusive for Western Ghats and consist of eight, six and one individuals respectively. 15 haplotypes of sambar were from all part of India and from overseas. The Neighbor-Joining (NJ) and Maximum Likelihood (ML) trees constructed on the basis of aligned sequences demonstrate four distinct cluster within the 21 haplotypes of sambar (Genus *Rusa*) and named as G_{1-4} (Figures 4.2 and 4.3). Major cluster of sambar was G_1, which appeared to be distributed from north to south of its range. Second cluster, G_2 consists of two sequences each of Javan sambar and sambar. Third cluster, G_3 shows interesting similarity of eastern sambar of India with that of the Malayan subspecies. The last cluster, G_4 was exclusive of Western Ghats and not observed in other part of India. Similar result has also obtained by Markov Chain Monte Carlo (*MCMC*) based Bayesian approach (Fig. 4.4). All the phylogenetic analyses suggest that G_4 appeared to be a unique cluster. Evaluation of sequence variation demonstrates an unambiguous 4% variation of G_4 with that of G_1, G_2 and G_3 (Table 4.1).

Result of different tree analysis (Figure 4.2-4.4) suggesting similar conclusions. Apart of above phylogenetic analysis, when genetic data of other cervides were compared with that of all sambar, an interesting phylogenetic position of Western Ghats (G_4) was observed. G_4 appeared to be one of the primitive clusters amongst examined sambar populations and Rusa species (Figure 4.5). Using genetic data, the age of the nodes was estimated and observed that select Western Ghats sambar (G_4) was evolve from its ancestor around 2.8 Mya, whereas, other sambar including Javan sambar (*R. timorensis*) were evolve around 1.2 Mya.

G_4 clade in all the phylogenetic trees (Figure 4.2-4.4) was strongly amalgamated with the out group; however, G_1, G_2 and G_3 were strongly supported with 99% bootstrap value. It genetically isolated the G_4 clade with the remaining three clades. Association of a recognized species *R. timorensis* (G_2) with the majority of the *R. unicolor* (G_1 and G_3) and differentiation of this cluster with the recognized *R. unicolor* (G_4) of Western Ghats highlighted an interesting phylogenetic disparity.

Figure 4.1: Different haplotypes (S_{1-21}) of genus *Rusa* and related sequence variations. Numerics after hyphen (_) are number of individuals of that haplotype. Dot (.) indicates similarity with that of first sequence.




220 230 240 250 260 270 280
S1_6 CCCATCAGACGCAGATAAAATTCCCTTCCACCCTTACTATACCATTAAAGATATCTTAGGCATCTTACTT
S2_6
S3_1
S4_1 T T C
S5_1 T T C
S6_2 T T C
S7_6 T T C
S8_51 T T T
S9_4 T T C
S10_2 T T T
S11_1 T T T T
S12_1 T T C
S13_1 T T
S14_1 T T
S15_1 T T
S16_1 T T
S17_1 T T
S18_1 T T T
S19_2 T T G
S20_1_RT T T G
S21_1_RT T T
Cervus_elaphus_FJ611891 C C T
290 300 310 320 330 340 350
S1_6 ATAGTACTCTTCTTAATATTGCTAGTATTATTCGCACCGGACCTGCTTGGAGACCCAGACAACTACACCC
S2_6
S3_1
S4_1 A C
S5_1 C T A
S6_2 C T A
S7_6 C T A
S8_51 C T A
S9_4 C T A
S10_2 C T A
S11_1 C T A
S12_1 C T A
S13_1 C T A
S14_1 A
S15_1 A
S16_1 A
S17_1 A
S18_1 C T A
S19_2 A A
S20_1_RT A A
S21_1_RT A A
Cervus_elaphus_FJ611891 C C A A T
360 370 380 390 400 410 420
S1_6 CAGCAAATCCACTCAACACACCCCCTCATATCAAACCCGAATGATATTTCCTATTTGCATACGCAATCCT
S2_6 T
S3_1
S4_1 T C
S5_1 T C
S6_2 T C
S7_6 T C
S8_51 T C
S9_4 T C
S10_2 T C
S11_1 T C
S12_1 T C
S13_1 T C
S14_1 T G C
S15_1 T G C
S16_1 T C
S17_1 T C
S18_1 T C
S19_2 T C
S20_1_RT T C
S21_1_RT T T C C
Cervus_elaphus_FJ611891 T C T T T

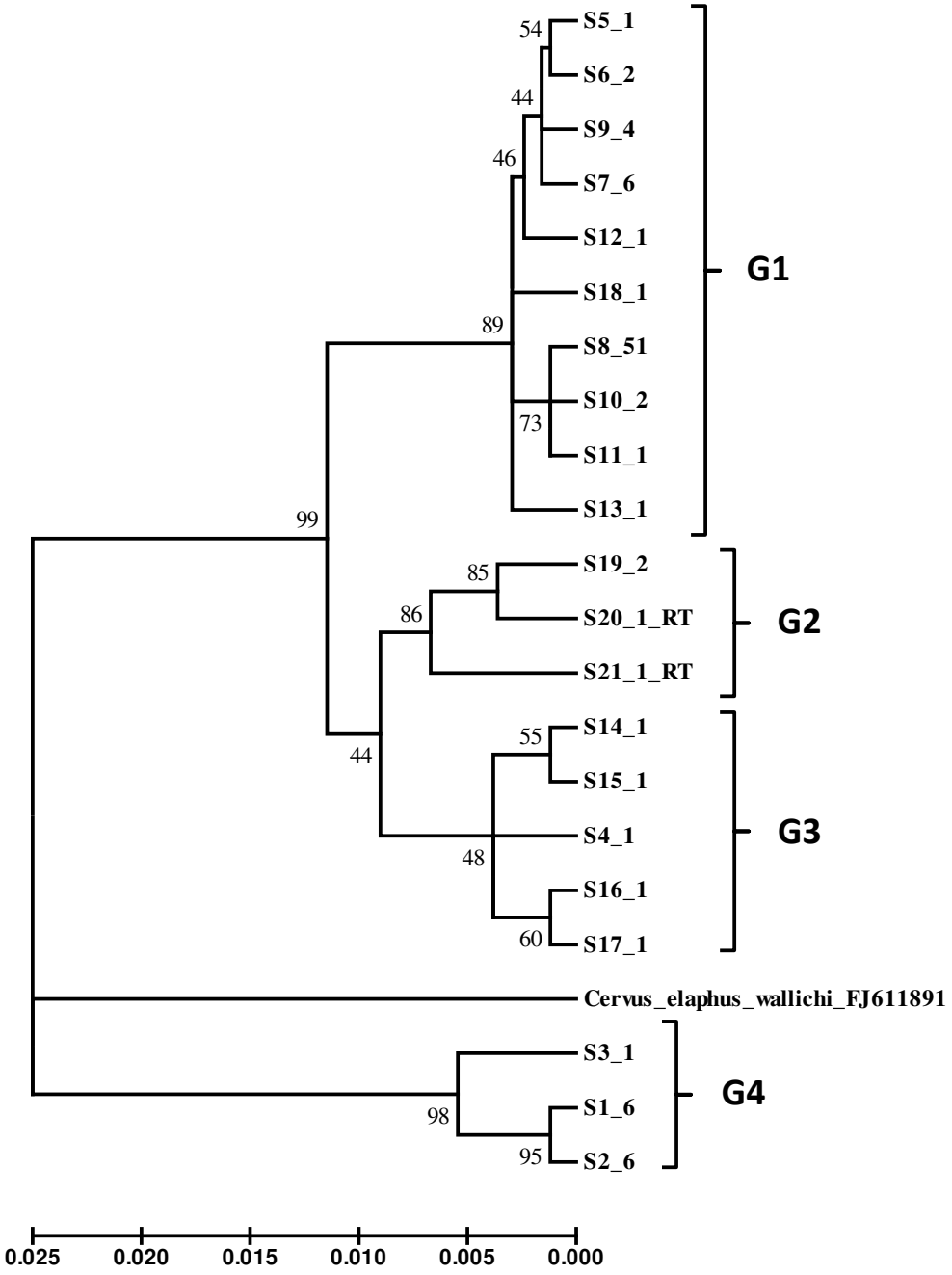


Figure 4.2: NJ tree showing the phylogenetic clustering among identified haplotype

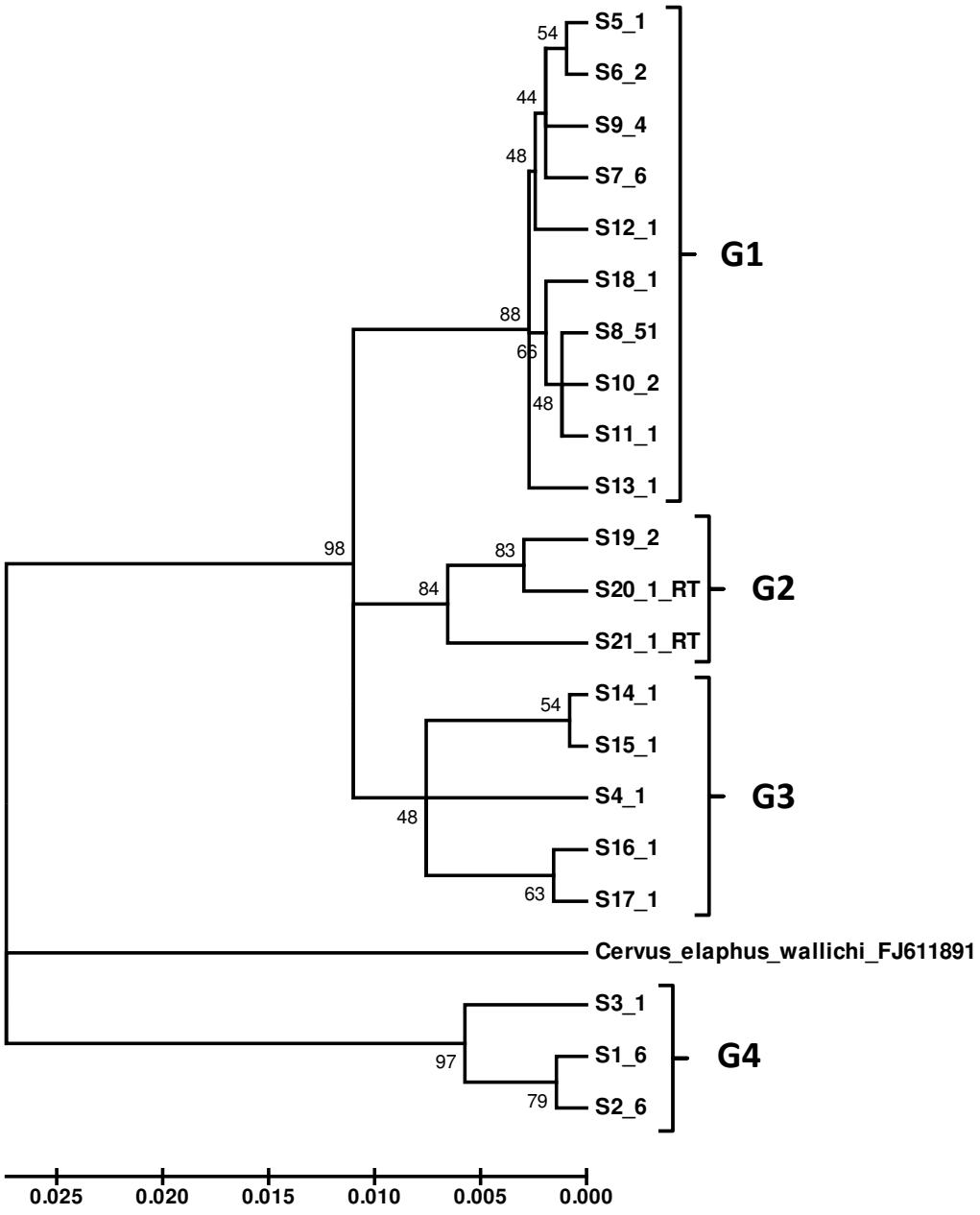


Figure 4.3: ML tree showing the phylogenetic clustering among identified haplotype

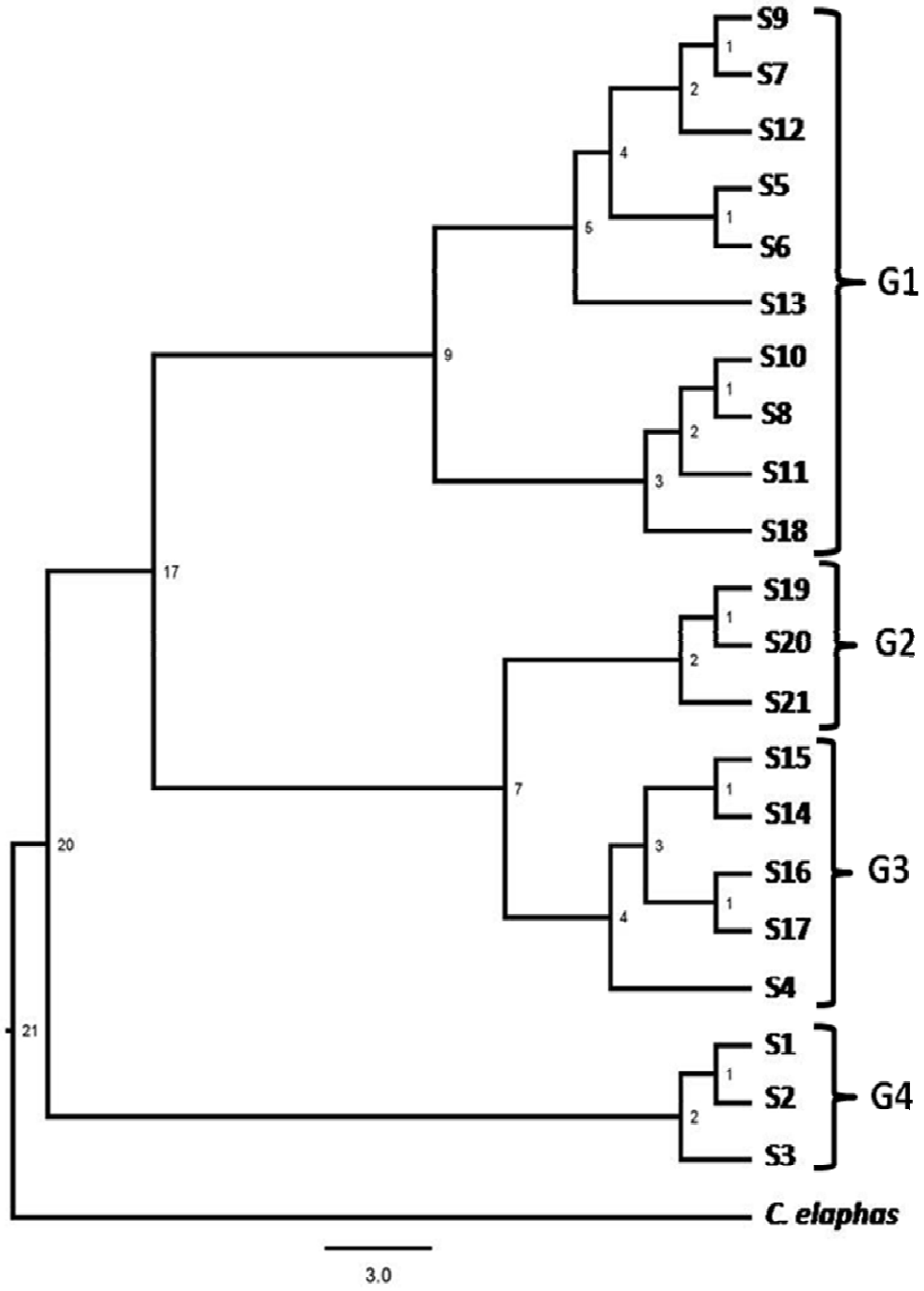


Figure 4.4: Markov Chain Monte Carlo (MCMC) based Bayesian tree showing the phylogenetic clustering among identified haplotypes of sambar

Table 4.1: Percentage similarity among identified haplotypes of *Rusa spp* and *Cervus elaphus*

	S7	S12	S8	S5	S6	S14	S18	S10	S11	S19	S20	S21	S9	S13	S16	S17	S15	S4	S1	S2	S3	C e
1: S7_6	**100.00**	99.76	99.29	99.52	99.76	98.57	98.81	99.05	99.05	97.62	97.39	97.86	99.76	99.76	98.34	98.10	98.34	98.10	95.72	95.49	95.72	94.77
2: S12_1	99.76	**100.00**	99.52	99.29	99.52	98.34	99.05	99.29	99.29	97.39	97.15	97.62	99.52	99.52	98.10	97.86	98.10	97.86	95.49	95.25	95.49	94.54
3: S8_51	99.29	99.52	**100.00**	98.81	99.05	98.34	99.52	99.76	99.76	97.39	97.15	97.62	99.05	99.52	98.10	97.86	98.10	97.86	95.49	95.25	95.49	94.54
4: S5_1	99.52	99.29	98.81	**100.00**	99.76	98.10	98.34	98.57	98.57	97.15	95.91	97.39	99.76	99.29	97.86	97.62	97.86	97.62	95.25	95.01	95.72	94.77
5: S6_2	99.76	99.52	99.05	99.76	**100.00**	98.34	98.57	98.81	98.81	97.39	97.15	97.62	99.52	99.52	98.10	97.86	98.10	97.86	95.49	95.25	95.49	94.54
6: S14_1	98.57	98.34	98.34	98.10	98.34	**100.00**	97.86	98.10	98.10	98.10	97.86	98.34	98.34	98.81	99.29	99.05	99.76	99.05	95.20	95.96	95.72	95.25
7: S18_1	98.81	99.05	99.52	98.34	98.57	97.86	**100.00**	99.29	99.29	97.39	97.15	97.62	98.57	99.05	98.10	98.34	98.10	97.86	95.49	95.25	95.49	94.54
8: S10_2	99.05	99.29	99.76	98.57	98.81	98.10	99.29	**100.00**	99.52	97.15	95.91	97.39	98.81	99.29	97.86	97.62	97.86	97.62	95.25	95.01	95.25	94.30
9: S11_1	99.05	99.29	99.76	98.57	98.81	98.10	99.29	99.52	**100.00**	97.15	95.91	97.39	98.81	99.29	97.86	97.62	97.86	97.62	95.25	95.01	95.25	94.30
10: S19_2	97.62	97.39	97.39	97.15	97.39	98.10	97.39	97.15	97.15	**100.00**	99.29	98.81	97.39	97.86	98.34	98.10	98.34	98.10	95.72	95.49	95.25	94.77
11: S20_1_RT	97.39	97.15	97.15	95.91	97.15	97.86	97.15	95.91	95.91	99.29	**100.00**	98.57	97.15	97.62	98.57	98.34	98.10	97.86	95.49	95.25	95.01	94.54
12: S21_1_RT	97.86	97.62	97.62	97.39	97.62	98.34	97.62	97.39	97.39	98.81	98.57	**100.00**	97.62	98.10	98.57	98.34	98.57	98.34	95.96	95.72	95.49	95.01
13: S9_4	99.76	99.52	99.05	99.76	99.52	98.34	98.57	98.81	98.81	97.39	97.15	97.62	**100.00**	99.52	98.10	97.86	98.10	97.86	95.49	95.25	95.96	95.01
14: S13_1	99.76	99.52	99.52	99.29	99.52	98.81	99.05	99.29	99.29	97.86	97.62	98.10	99.52	**100.00**	98.57	98.34	98.57	98.34	95.96	95.72	95.96	95.01
15: S16_1	98.34	98.10	98.10	97.86	98.10	99.29	98.10	97.86	97.86	98.34	98.57	98.57	98.10	98.57	**100.00**	99.76	99.52	99.29	95.44	95.20	95.96	95.49
16: S17_1	98.10	97.86	97.86	97.62	97.86	99.05	98.34	97.62	97.62	98.10	98.34	98.34	97.86	98.34	99.76	**100.00**	99.29	99.05	95.20	95.96	95.72	95.25
17: S15_1	98.34	98.10	98.10	97.86	98.10	99.76	98.10	97.86	97.86	98.34	98.10	98.57	98.10	98.57	99.52	99.29	**100.00**	99.29	95.44	95.20	95.96	95.49
18: S4_1	98.10	97.86	97.86	97.62	97.86	99.05	97.86	97.62	97.62	98.10	97.86	98.34	97.86	98.34	99.29	99.05	99.29	**100.00**	95.20	95.96	95.72	95.25
19: S1_6	95.72	95.49	95.49	95.25	95.49	95.20	95.49	95.25	95.25	95.72	95.49	95.96	95.49	95.96	95.44	95.20	95.44	95.20	**100.00**	99.76	99.05	95.25
20: S2_6	95.49	95.25	95.25	95.01	95.25	95.96	95.25	95.01	95.01	95.49	95.25	95.72	95.25	95.72	95.20	95.96	95.20	95.96	99.76	**100.00**	98.81	95.49
21: S3_1	95.72	95.49	95.49	95.72	95.49	95.72	95.49	95.25	95.25	95.25	95.01	95.49	95.96	95.96	95.96	95.72	95.96	95.72	99.05	98.81	**100.00**	95.25
22: C_elaphus	94.77	94.54	94.54	94.77	94.54	95.25	94.54	94.30	94.30	94.77	94.54	95.01	95.01	95.01	95.49	95.25	95.49	95.25	95.25	95.49	95.25	**100.00**

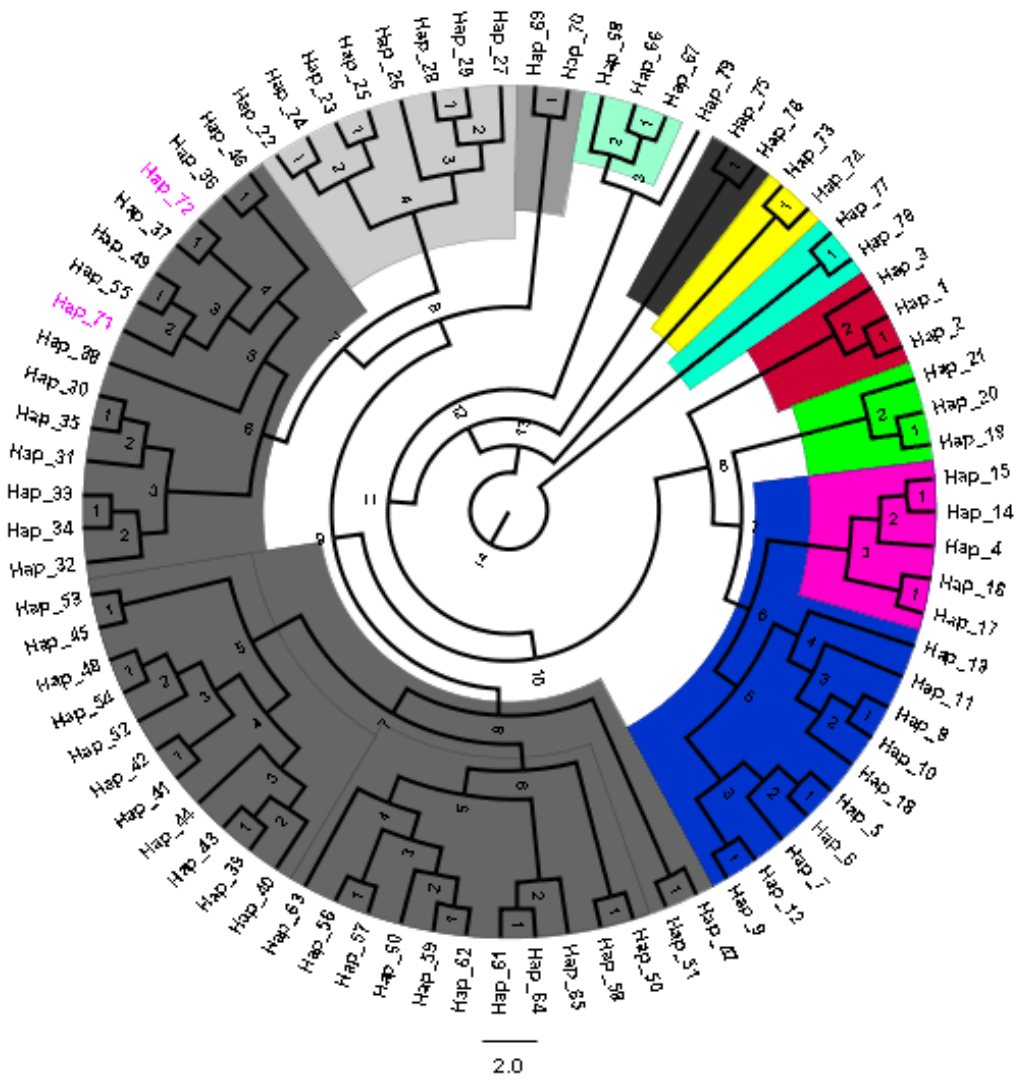


Figure 4.5: Markov Chain Monte Carlo (*MCMC*) based Bayesian approach showing the phylogenetic clustering among identified haplotypes of sambar with that of other cervides

4.4 DISCUSSION

Phylogeography is a well established tool for detection of distinct lineage and centre for evolution of a species (Larson *et al*. 2004). During a phylogenetic study on old world deer it was suggested that fellow deer (*Dama dama*) evolved 5.2 million year ago (mya) (Pitra *et al*. 2005). Based on evolutionary clock derived from genetic data, this study detected an ancient lineage of genus *Rusa* in Western Ghats, which evolved 2.8 mya. The present *R. unicolor* and *R. timorensis* evolved around one mya in south-east Asia (near Javan Island). Paleontology also support our data that *R. timorensis* apparently did not appear until the late Pleistocene (Dong 1993). A larger size fossil *R. unicolor* has been found in caves in northern Vietnam, which could be from the middle Pleistocene 0.8-1.7 mya (Bacon et al. 2004). It indicates that modern *R. unicolor* is a recent species between 0.8-1.7 my (Bacon *et al*. 2004; Huang *et al*. 1995; Si *et al*. 1993). Furthermore, it was evidenced from the analysis that *R. timorensis* and sambar of Western Ghats were evolved from a common ancestor.

Existence of two close groups at very distinct locations (one in Western Ghats and another in Javan Island) can be explained as probably (1) their common ancestor was widely distributed in Asia around 2.8 mya with the progress of Pleistocene epoch after 2.5 mya (Gribbin 1982) they were probably fragmented and restricted into two populations, one in Western Ghats as it is well known refugia for funna (Prasad *et al*. 2009) and another in Javan Island. (2) their movement in western Ghats could be historic and they might be evolved at Javan region and migrated towards peninsular India via north-east rout (Mani 1974) and harsh condition of Pleistocene could be responsible for wiping out of the most of the population except refugia population of Western Ghats of India. When climatic conditions become favorable either (1) the south-east Asian ancestor tried to expand its range in remaining favorable habitats of south Asia or (2) Western Ghats population started radiating towards northern part and subsequently entered into south-and south-east Asia through Assam gateway. During this process, around 1.2 mya two recognized and surviving species of *Rusa* were evolved i.e. *R. unicolor* and *R. timorensis.* Similar observation has also been derived from fossil records (Leslie 2009; Schaller 1998). Moreover, as

observed that sambar in Western Ghats were older (2.8 mya) than remaining populations (1.2 mya), similarly Mani (1974) suggested that peninsular autochthonous elements are much older age than the radiated elements of south-east Asia. Hence the present study complements the inference of fossil and biogeographical observations.

Morphologically *R. unicolor, R. timorensis* and Western Ghats sambar are not much distinct except they have minor variations. The cyt *b* gene based analysis indicated that due to insufficient morphological variation two sequences of *R. timorensis* (FJ556574-75) were mistaken submitted as *R. unicolor* from Thailand. Another probability of mistaken identification could be these two samples would be of hybrid individuals with probable maternal lineage of *R. timorensis*. It indicated that the *Rusa spp*. do not have unambiguous morphological variations among the recognized species, which leads to mistaken identification of them. However they have sufficient genetic variation in cyt *b* gene with minimum of 4% difference (Table 4.1). Molecular phylogeny is an unambiguous approach for evaluation of the phylogenic structure of the species, especially; when morphological feature is not clearly visible among the close species. Phylogenic uncertainty in conventional taxonomy has been resolved using cyt *b* gene (Pitra et al. 2004). Despite of many debates on use of accurate molecular marker in phylogenetic study, cyt *b* proven a better marker in compared to other mtDNA markers, which provides more nucleotide variation in shorter fragment (Tobe *et al*. 2010). Present *R. timorensis*, appeared to be recent descendent of this cluster, which has around 4% variation in cyt *b* gene with that of Western Ghats lineage as well other *R. unicolor*. In recent past, genetics has been widely accepted as a supportive tool in phylogenetic revisions (Tirosh *et al.* 2006; Landry *et al.* 2007). Populations that show substantial genetic differences may be classified as different species (Price 2008).

Cytogenetically, two nascent species of rusine group *R. unicolor* and *R. timorensis* were distinct. The *R. unicolor* has 62 chromosomes and *R. timorensis* has 60 chromosomes (Idris and Moin 2009). Reproductive isolation in nascent species of rusine group was not observed and it was recorded that

R. unicolor and *R. timorensis* could bred naturally and able to produced fertile offspring (Idris and Moin 2009). Hence, in nascent species reproductive isolation should not be a barrier to be considered as distinct species. Recent phylogenetic revision in rusine deer indicated that this modern group of the cervides is highly fragile. Taxonomic alteration in this group resulted into the separation of one species *R. marianaa* from the subspecies of *R. unicolor* (Grubb 2005). However, this taxonomic revision is still a debatable topic (Francis 2009). Similarly, based on this study it was appeared that sambar of Western Ghats could probably an ecological significant unit (ESU), which is older among all examined rusa samples. Interesting taxonomic revision has been witnessed in last two decades for sambar (Grubb 1990). Researchers indicated existence of novel variation in large deer in Indochina region and suspected that a new species might be involved (Anh *et al.* 1996). Based on previous taxonomic changes, additional taxonomic revision among sambar was warranted (Timmin *et al.* 2008). However, before drawing a firm conclusion for the Western Ghats lineage a deep ecological study is required assess the existence of significant variations, if any.

Based on the preliminary information obtained from this Ph.D. work, a pilot study was also conducted on a sympatric ungulate species wild pig, *Sus scrofa* (Gupta *et al.* 2013b). It was focused to screen the genetic variation among wild pigs in India and to draw its phylogenetic position. The Cyt *b* sequences of the 42 Indian wild pig samples exhibited only 97% homology with those of the domestic pigs and other wild pig races of the world. The 3% genetic variation within one of the evolutionary conserved genes between the Indian wild and domestic pigs and the other wild pig races separated the Indian wild pigs from the other two forms (Gupta *et al.* 2013b). It was apparent from molecular data that domestic pigs in India were not descendent of the main Indian wild pig, however; they exhibited a close affinity with the other wild pig races found in remaining part of Asia and Europe (Gupta *et al*. 2013b). It was indicated from these two examples (sambar and wild pigs) that similar phylogenetic disparity might also be existing in other species, which occurs in India along with other neighboring south-east Asian countries.

4.5 CONCLUSIONS

This study indicates that the sambar population in Western Ghats is ancient lineage, which probably did not participated in active evolution process and refrain wide distribution so as to restrict themselves in the vicinity of Western Ghats. The Javan ancestor further evolved into two distinct species of genus Rusa i.e. *R. unicolor* and *R. timorensis.* It was assumed that south-east Asia could be a centre for speciation of modern *R. unicolor* (around 1 mya). More sampling from this region is required to reach a firm conclusion. However, this study appeared to be indicating that the Western Ghats lineage is genetically distinct from remaining *R. unicolor* and *R. timorensis.* The select Western Ghats sambar appeared to be ancient among extant rusa and could probably be an evolutionary significant unit, which has significant older divergence time of 2.8 mya than the remaining *R. unicolor* and *R. timorensis* (around 1 mya).

CHAPTER 5
GENETIC DIVERSITY AND GENE FLOW

SUMMARY

Based on genotyping data, the genetic variability among sambar populations was examined for measuring the level of heterozygosity. The mean number of alleles per locus was 10.74 and mean proportion of individuals typed 0.58. The mean polymorphic information content (PIC) value was 0.711, which indicated that tested loci were highly polymorphic for sambar species. The mean expected and observed heterozygosity was 0.752 and 0.528, respectively. The observed heterozygosity is significantly lower, which indicates significant genetic disparity from expected Hardy-Weinberg saturation. This study suggests that all the *Rusa spp*. were perhaps evolved from common ancestor and exclusive population of Western Ghats was primitive among them. It was evolved around 2.8 mya which probably lost its connectivity with that south-east Asian population during unfavorable conditions of ice-age (around 2.5 mya). Further, upon normalization of climatic conditions at the end of ice-age the south-east Asian counterpart started radiating, and speciation in to *R. unicolor* and *R timorensis* took place at around 1.2 mya. The recent sambar populations occupying in Thailand, Malayan and Indian region could be a result of radiation of the recent clade of south-east Asia. Whereas, the Western Ghats clade could probably be ancient and surviving populations of this group.

5.1 BACKGROUND

Variation in alleles of a locus is known as genetic variation, which can be observed both within and among populations. Genetic variation is both a trait of individuals and a trait of populations (Lacy 1997). Variation within individuals of diploid species is most commonly characterized by the percentage of loci at which an individual is heterozygous. Variation within populations also includes inter-individual variation, and often is quantified by the gene diversity (the heterozygosity expected under Hardy-Weinberg Equilibrium), by the number of distinct alleles per locus, or by the percentage of loci that are polymorphic (Nei 1973). Mean within-individual variation usually is highly correlated with populational (between-individual) variation and all measures of populational variation encompass within individual variation in

the population. Heterozygosity is depleted by inbreeding (mating between relatives), which leads to a greater probability of the two alleles at a locus being identical by descent from an ancestor common to both sides of the pedigree, and by genetic drift (Lacy 1997). Genetic variation is important because it provides the genetic material for natural selection. Genetic variation is brought about by mutation, which is a permanent change in the chemical structure of a gene. The mutation is started off by a parent, as the parent mates the offspring now has a chance to receive that mutation trait also. Now when that mutated offspring is ready to mate they now have the chance of passing on that trait to their offspring. This process begins the first generation of mutated offspring.

Genetic variation within a population is commonly measured as the percentage of gene loci that are polymorphic or the percentage of gene loci in individuals that are heterozygous. A variety of factors maintain genetic variation in populations. Potentially harmful recessive alleles can be hidden from selection in the heterozygous individuals in populations of diploid organisms (recessive alleles are only expressed in the less common homozygous individuals). Natural selection can also maintain genetic variation in balanced polymorphisms. Balanced polymorphisms may occur when heterozygotes are favored or when selection is frequency dependent. Population level genetic variation was examined by sequencing mtDNA control region and genotyping of STR loci. Sequencing data of mtDNA loci were also used in estimation genetic diversity and prediction of gene flow. During the field study biological samples were collected across India. The majority of the samples were from central highland (Panna Tiger Reserve) and Western Ghats (Goa, Tamil Nadu and Kerala States). 23 deer specific STR primers (Bonnet *et al.* 2002, Gaur *et al.* 2003) were used for evaluation of genetic fitness of these two populations. The forward (F) primers of each STR loci were labelled with fluorescent dye (Table 2.4).

5.2 MATERIAL AND METHODS

5.2.1 DNA extraction and amplification (same as Section 4.3.1)

The extracted DNA was used for multiplex PCR amplification with the primers described in table 2.4. PCR reactions were carried out in 10µl reaction volume using 1x Qiagen Multiplex PCR kit, 4 pmol of each primers and 1µl (~30 ng) template DNA. The PCR conditions were: initial denaturation at 95°C for 15 min, followed by

35 cycles of denaturation at 95°C for 45s, annealing at 54°C for 40s, and extension at 72°C for 75s. The final extension was at 72°C for 10 min.

5.2.2 Post PCR cleanup, DNA sequencing

The 1.5 µl of each PCR product was incubated with 0.5 µl ExoSAP-IT (USB, Cleveland, Ohio) at 37°C for 20 min. followed by inactivation of enzymes at 85°C for 15 min. The PCR products after *Exo-I* and *SAP* treatment were sequenced using ABI 3130 Genetic Analyzer on both strands. The sequences generated for both the strands were examined and resolved for any ambiguity by using Sequencher 4.1 programme.

5.2.3 Genotyping

All the PCR products amplified by fluorescent labeled dye were used for genotyping in ABI 3130 along with ABI size standard labeled with Liz dye. Base pair calling was done using GeneMapper v. 3.7 software (ABI, *Life Technology*). The genotyping data obtained from the tested samples were used for further genetic analysis.

5.2.4 Genotyping data analysis

5.2.4.1 Estimation of genetic diversity and Hardy-Weinberg equilibrium

The observed (H_O) and expected (H_E) heterozygosity and polymorphic information content (*PIC*) value were computed using *CERVUS 3.0* (Kalinowski et al 2007). Deviation from Hardy-Weinberg equilibrium (*HWE*) was estimated using GENEPOP 4.2.2 software (Raymond and Rosset 1995).

5.2.5 Bayesian analysis

5.2.5.1 Structuring the population using STRUCTURE software

The Bayesian analysis for clustering and inferring populations of sambar was carried out using STRUCTURE 2.3.3 (Pritchard *et al.* 2000). The clustering algorithm utilized is meant to identify distinctive clusters and also to infer how these clusters relate to geographical data of individuals of these populations. The distance based genetic methods are usually dependent on both the distance measures and graphical representation chosen. Moreover, the confidence in this clustering pattern is low. Thus, the distance based method is suited for exploratory data analysis than to find the statistical inference. The structure software implements a model based approach. For

the continuity of sambar populations (there being no strict physical barriers or isolation of the populations), the admixture model was employed. In data analysis 15 independent runs of Gibbs sampler for each of K (K= 2 to 10) were run. The results presented are based on a burnin value of 60,000 followed by recording 1,00,000 MCMC (Monte Carlo Marcov Chain) simulations.

Fifteen independent replicates at each K value were run. The appropriate K value was determined by calculating ad hoc quantity (ΔK) as proposed by Evanno *et al*. (2005) using STRUCTURE HARVESTER software (Earl and vonHoldt 2012). Each individual was assigned to the inferred clusters using a threshold proportion of membership (q), i.e. $q \geq 0.80$, otherwise an individual was determined as admixed if the q-value was less than 0.80.

For structure analysis, multilocus genotyping data of 49 individuals was used (Table 5.1). These individuals were produced genotype data of more than 9 loci out of 23 tested loci.

Table 5.1: List of samples used for Bayesian structure analysis

S. No.	Sampling location	Indian State	No. of individuals
1	Miscellaneous	Uttarakhand, Himachal Pradesh, Rajasthan, and Chhattisgarh	10 (1-10)
2	Deccan Peninsula	Panna Tiger Reserve, M.P.	16 (11-26)
3	Western Ghats	Goa, Kerala and Tamil Nadu	23 (27-49)

5.2.6 Evaluation of gene flow

Prediction of historic movement of various populations (subspecies) of a species provides an insight into the probable source of origin of that species. Sequence data generated from different population can be used for establishment of their historic connectivity. The Network 4.612 program was used to analyze the evolutionary relatedness based on the cyt *b* gene sequences of sambar.

5.3 RESULTS

5.3.1 Genetic diversity among sambar populations

Out of 23 loci, only two loci (INRABERN185 and ILSTS005) were monomorphic in Deccan Peninsula population. Genetic diversity for 23 loci of two biogeographic zones population was estimated for comparison of genetic variability among them. Genetic diversity of the biogeographic zones wise populations were as:

5.3.1.1 Genetic diversity of Deccan Peninsula populations

Based on genotyping data, the genetic variability was estimated for measuring the level of heterozygosity of Deccan Peninsula (Central highland) populations (Table 5.2). The mean number of alleles per locus was 6.65 and mean proportion of individuals typed was 0.556. It showed the markers used were appropriate to assessment of genetic fitness in given sambar populations. Two loci INRABERN185 and ILSTS005 were monomorphic for tested central Indian individuals (Table 5.2). The mean polymorphic information content (PIC) value was 0.642 indicating that tested loci are highly polymorphic for sambar species. The mean expected and observed heterozygosity was 0.713 and 0.474, respectively. The observed heterozygosity was significantly lower, which was an indication for significant genetic disparity from expected Hardy-Weinberg saturation.

5.3.1.2 Genetic diversity of Western Ghats populations

Based on genotyping data, the genetic variability was estimated for measuring the level of heterozygosity of Western Ghats populations (Table 5.3). The mean number of alleles per locus was 6.65 and mean proportion of individuals typed was 0.56. It showed the markers used are appropriate to assessment of genetic fitness in given sambar populations. All the loci were polymorphic for tested Western Ghats individuals**.** The mean polymorphic information content (PIC) value was 0.653 indicating that tested loci are highly polymorphic for sambar species. The mean expected and observed heterozygosity was 0.720 and 0.533, respectively. Most of the loci yielded comparable expected and observed heterozygosity (Table 5.3). The observed heterozygosity was slightly lower, which was an indication for genetically healthier than that of Deccan Peninsular population.

5.3.1.3 Overall genetic diversity

Based on Multilocus genotyping data obtained from two different biogeographic zone populations the observed mean number of alleles per locus was 10.74 and mean proportion of individuals typed was 0.58 (Table 5.4). It showed the markers used were appropriate to assessment of genetic fitness in sambar populations. The mean polymorphic information content (PIC) value was also significantly higher (0.711), indicated that tested loci were highly polymorphic for sambar species. The mean expected and observed heterozygosity was 0.752 and 0.528, respectively. The observed heterozygosity was significantly lower, which is an indicating for significant genetic disparity from expected Hardy-Weinberg saturation.

Table 5.2: Summary of diversity statistics of Deccan Peninsula population

S. No.	Locus	Allele (N)	Range (bp)	Heterozygosity		*PIC*	HW	F(Null)
				H_O	H_E			
1	BM4208	7	145-159	0.563	0.782	0.723	ND	+0.1453
2	INRA011	8	198-220	0.533	0.825	0.773	ND	+0.2073
3	Ca18	9	164-222	0.643	0.897	0.850	ND	+0.1460
4	CSSM1	9	152-188	0.429	0.849	0.798	ND	+0.3247
5	RT1	8	204-222	0.429	0.807	0.749	ND	+0.3070
6	Cervid1	6	162-186	0.125	0.742	0.666	NS	ND
7	BM4107	13	140-174	0.615	0.923	0.878	ND	+0.1827
8	T193	4	174-208	0.400	0.533	0.450	ND	ND
9	T108	6	132-158	0.500	0.617	0.553	ND	ND
10	NVHRT48	8	74-128	0.545	0.766	0.703	NS	+0.1772
11	CELJP27	8	161-189	0.333	0.732	0.676	NS	+0.3710
12	D	5	138-184	0.000	0.810	0.728	ND	ND
13	BM6506	9	183-211	0.467	0.867	0.819	ND	+0.2769
14	RT27	6	145-177	1.000	0.929	0.786	ND	ND
15	INRABERN185	1	238	0.000	0.000	0.000	ND	ND
16	ILSTS005	1	165	0.000	0.000	0.000	ND	ND
17	T123	3	155-163	0.400	0.378	0.314	ND	ND
18	T507	6	142-164	0.500	0.850	0.768	ND	ND
19	T156	8	113-189	0.833	0.894	0.800	ND	ND
20	OarFCB193	14	82-134	0.538	0.929	0.885	ND	+0.2427
21	NVHRT16	3	164-200	0.833	0.591	0.460	ND	ND
22	RT6	7	87-117	0.571	0.813	0.730	ND	ND
23	Ca42	4	177-195	0.667	0.867	0.671	ND	ND
		6.65		**0.474**	**0.713**	**0.642**		

H_O **-** Observed heterozygosity**;** ***H_E*** **-** Expected heterozygosity; ***PIC***Polymorphic information content; ***HWE*** **-** Hardy Weinberg Equilibrium; **F(null) -** Frequency of predicted null alleles

Table 5.3: Summary of diversity statistics of Western Ghats population

S. No.	Locus	Allele (N)	Range (bp)	Heterozygosity		*PIC*	HW	F(Null)
				H_O	H_E			
1	BM4208	8	137-161	0.773	0.747	0.689	NS	-0.0489
2	INRA011	8	192-222	0.727	0.746	0.700	NS	+0.0065
3	Ca18	8	164-210	0.478	0.709	0.664	NS	+0.1909
4	CSSM1	8	154-210	0.579	0.788	0.737	NS	+0.1485
5	RT1	4	202-208	0.455	0.635	0.547	NS	+0.1603
6	Cervid1	4	156-166	0.294	0.273	0.253	ND	-0.0714
7	BM4107	5	144-152	0.533	0.777	0.712	NS	+0.1668
8	T193	9	168-208	0.750	0.847	0.799	ND	+0.0441
9	T108	9	132-162	0.611	0.802	0.752	NS	+0.1082
10	NVHRT48	10	74-120	0.722	0.849	0.804	ND	+0.0726
11	CELJP27	3	185-191	0.294	0.358	0.320	ND	+0.0669
12	D	8	138-184	0.385	0.729	0.675	NS	+0.2675
13	BM6506	10	183-205	0.333	0.824	0.780	NS	+0.4186
14	RT27	3	163-175	0.333	0.733	0.535	ND	ND
15	INRABERN185	5	228-246	0.125	0.792	0.701	ND	ND
16	ILSTS005	4	163-177	0.111	0.542	0.473	ND	ND
17	T123	6	135-165	0.786	0.714	0.637	NS	-0.0923
18	T507	4	144-164	0.692	0.760	0.681	ND	+0.0217
19	T156	9	131-181	0.727	0.900	0.844	ND	+0.0772
20	OarFCB193	10	86-136	0.733	0.862	0.815	ND	+0.0698
21	NVHRT16	3	164-188	0.273	0.498	0.419	ND	+0.2521
22	RT6	6	91-113	0.750	0.775	0.691	ND	ND
23	Ca42	7	175-199	0.800	0.911	0.798	ND	ND
		6.65		**0.533**	**0.720**	**0.653**		

H_O **-** Observed heterozygosity; ***H_E*** **-** Expected heterozygosity; ***PIC*** **-** Polymorphic information content; ***HWE*** **-** Hardy Weinberg Equilibrium**; F(null) -** Frequency of predicted null alleles

Table 5.4: Summary of diversity statistics of all tested samples

S. No.	Locus	Allele (N)	Range (bp)	Heterozygosity		*PIC*	HW	F(Null)
				H_O	H_E			
1	BM4208	12	137-163	0.688	0.747	0.703	NS	+0.0205
2	INRA011	12	152-222	0.638	0.757	0.724	NS	+0.0837
3	Ca18	14	154-222	0.574	0.819	0.790	NS	+0.1723
4	CSSM1	16	152-210	0.595	0.885	0.862	ND	+0.1856
5	RT1	9	202-222	0.452	0.678	0.616	NS	+0.2050
6	Cervid1	9	156-186	0.281	0.480	0.461	NS	+0.2470
7	BM4107	13	140-174	0.531	0.833	0.799	NS	+0.2128
8	T193	12	168-212	0.680	0.849	0.814	ND	+0.1085
9	T108	11	124-162	0.548	0.749	0.711	NS	+0.1337
10	NVHRT48	13	74-128	0.676	0.817	0.781	NS	+0.0962
11	CELJP27	9	161-191	0.294	0.522	0.489	NS	+0.2785
12	D	10	138-184	0.231	0.750	0.711	NS	+0.5114
13	BM6506	13	173-211	0.421	0.865	0.839	NS	+0.3399
14	RT27	9	145-177	0.750	0.883	0.810	ND	ND
15	INRABERN185	5	228-246	0.091	0.723	0.643	NS	+0.7753
16	ILSTS005	4	163-177	0.067	0.356	0.324	ND	+0.7055
17	T123	8	135-167	0.700	0.676	0.621	NS	-0.0625
18	T507	6	142-164	0.636	0.789	0.733	ND	+0.0840
19	T156	14	113-189	0.778	0.903	0.867	ND	+0.0611
20	OarFCB193	20	82-136	0.688	0.920	0.899	ND	+0.1295
21	NVHRT16	6	164-200	0.571	0.613	0.564	NS	-0.0010
22	RT6	11	87-117	0.625	0.770	0.730	NS	+0.0825
23	Ca42	11	169-199	0.636	0.926	0.873	ND	+0.1608
		10.74		**0.528**	**0.752**	**0.711**		

H_O **-** Observed heterozygosity, ***H_E*** **-** Expected heterozygosity, ***PIC*** **-** Polymorphic information content; ***HWE*** **-** Hardy Weinberg Equilibrium; **F(null) -** Frequency of predicted null alleles

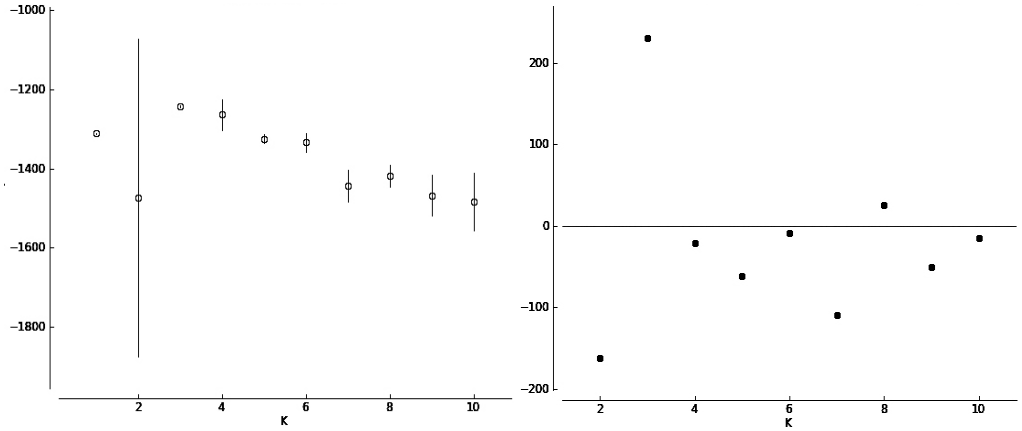
5.3.2 Population genetic structure

The log-likelihood [L(K)] values and ad hoc quantity (ΔK) are given in (Table 5.5 and Figure 5.1A). The ad hoc quantity (ΔK) was used to determining the real K value (K=3 for 49 individuals) (Table 5.5 and Figure 5.1 D). Successive increase in the K values did not split the three major groups into additional clusters. Hence, three distinguishable populations from the genotyping data were identified using Structure program. The individuals were assigned to one or more clusters on the basis of their q values, if the q value was equal or larger than the probability threshold of 0.800 then the individuals were assigned to a single cluster while individuals with q value lower than 0.800, were assigned to more than one clusters (admixture). The random clustering of populations has been represented in figure 5.2.

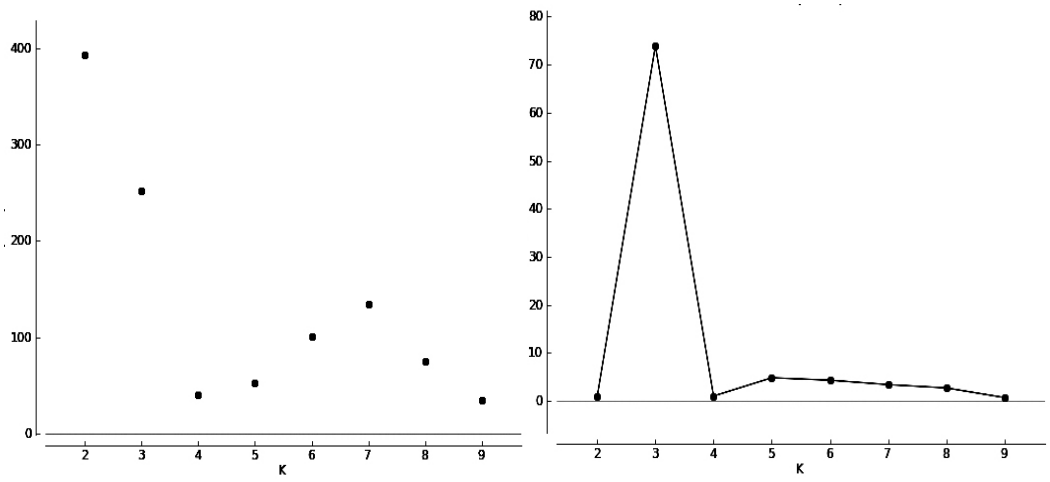
From structure analysis it was evident that one population was invariably restricted in only Western Ghats as indicated by blue bar shown in K=3 and in dotted line at remaining K values (Figure 5.2). Population pattern was unchanged for these selected Western Ghats individuals, it was conferred after comparison of all bar plot between K=3 to 10 (dotted line in figure 5.2). However, one admix individual with below the significant contribution (< 0.8) from this blue population was also observed in Deccan Peninsular region indicating the probable gene flow from this population to other area. Remaining two populations was uniformly distributed in rest of the study area including Western Ghats, Deccan Peninsula and other tested biogeographic zones.

Table 5.5: Inferring the value of K, the number of populations, for all tested sambar populations

No. of assumed populations (K)	Mean L (K)	Ln'(K)	Ln''(K)	Δ K
1	-1311.38	0.00	0.00	0.00
2	-1473.62	-162.24	392.98	0.978057
3	-1242.88	230.74	251.78	**73.91878**
4	-1263.92	-21.04	39.92	1.001828
5	-1324.88	-60.96	52.1	4.800528
6	-1333.74	-8.86	100.44	4.303609
7	-1443.04	-109.3	134.62	3.383616
8	-1417.72	25.32	75.22	2.695976
9	-1467.62	-49.9	34.72	0.683985
10	-1482.80	-15.18	0.00	0.00



A-Mean L(K) **B-**L´(K)= L(K)-L(K-1)



C-L´´(K)=L´(K+1)- L´(K) **D-** ΔK= L´´(K)/ Stdv L(K)

Figure 5.1: Graphical representation of the true number of groups

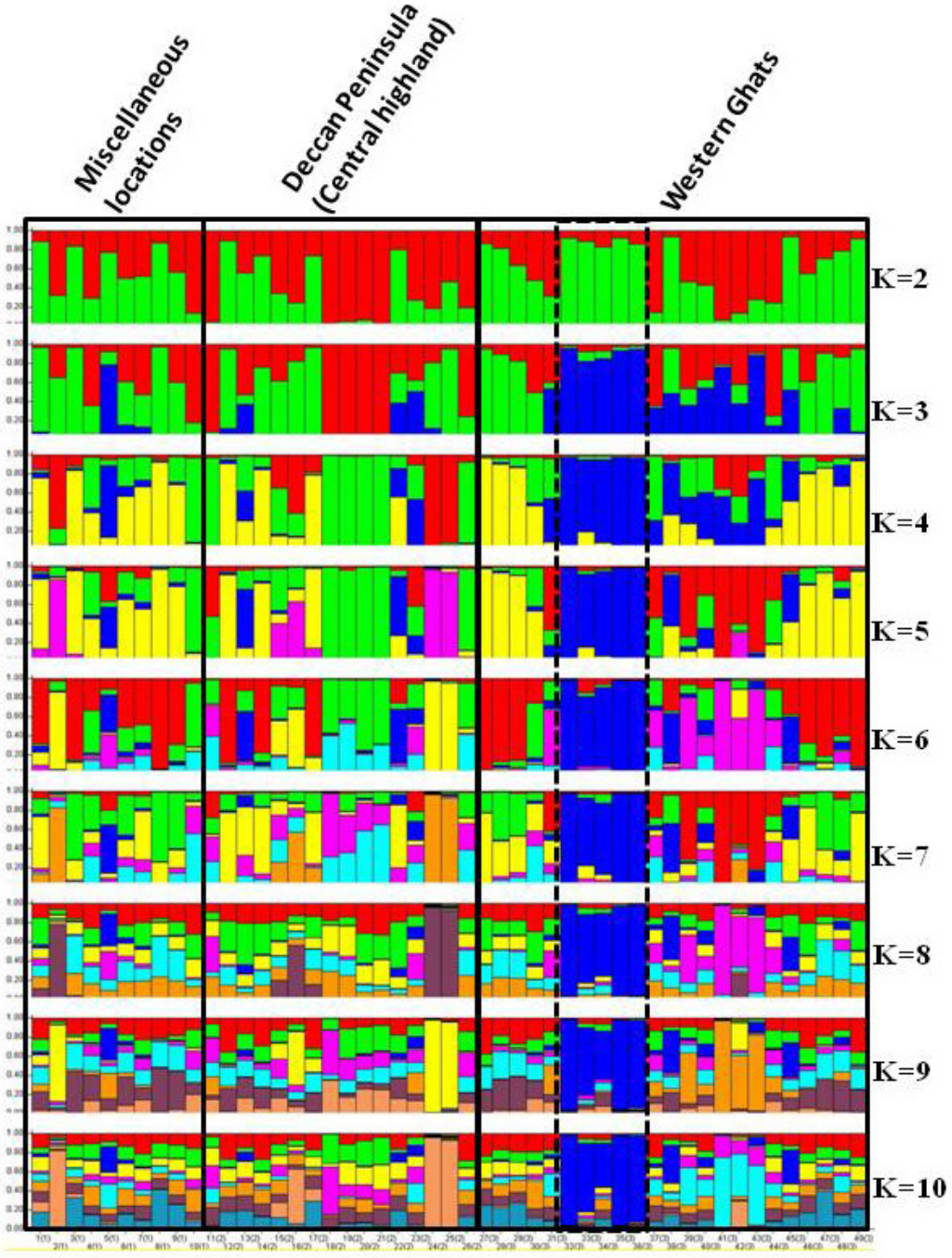


Figure 5.2: Random assignment of the individuals to population using Structure 2.3.4 program at different K values (2-10). Individuals of Western Ghats population in dotted line indicating no change in population assignment for these individuals at different K values.

5.3.3 Evolution of *Rusa spp.* and gene flow

Upon comparison of cyt *b* sequence of Indian and overseas sambar along with the Javan rusa using NETWORK software, three distinct networking cluster was observed among the Rusa species (Figure 5.3). First cluster was of Western Ghats sambar (H_1-H_3) and Javan Rusa. Second cluster was of Malayan, Thailand (H_{14}-H_{16}) and eastern India (H_5 and H_{13}). Third cluster was of north, central and south India (H_6-H_{12}). First cluster appeared to indicating that the Javan sambar and sambar of Western Ghats originated from a common ancestor. The remaining sambar population of the world appeared to be originated from a common ancestor.

When the sequences of other cervides species (H_{22}-H_{76}) were included in networking analysis, networking of all Rusa species was observed in one cluster and rest of the cervides were in a distinct cluster (Figure 5.4). Among the Rusa cluster the first node (indicated by arrow in figure 5.4) was formed for exclusive populations of Western Ghats (H_1-H_3). It indicates that this node could be a primitive among the Rusa species. Subsequent to this primitive node, it was further divided in to two group one was of Javan Rusa (H_{19}-H_{20}) and another was of remaining sambar populations of India (H_4-H_{13}), Malayan and Thialand (H_{14}-H_{18}).

The above analysis appeared to be indicating that ancient divergence of Rusa group was occurred before the speciation of *R. unicolor*. The Western Ghats sambar and Javan Rusa cluster was separated much earlier before the speciation of *R. unicolor* from its ancestor. The major group of *R. unicolor* has radiated throughout its distribution range till the southern boundary of India. On other hand Western Ghats exclusive sambar was appeared to be confined to the vicinity of it and did not radiated much.

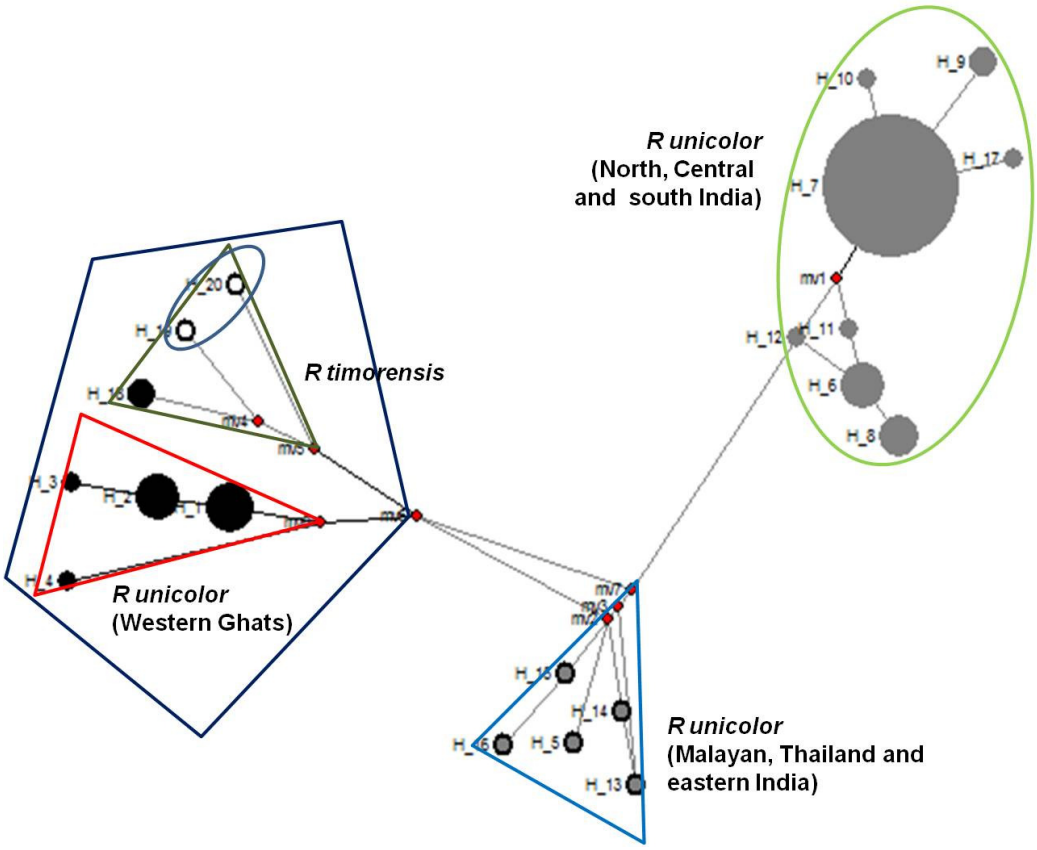


Figure 5.3: Networking of *Rusa species* of world.

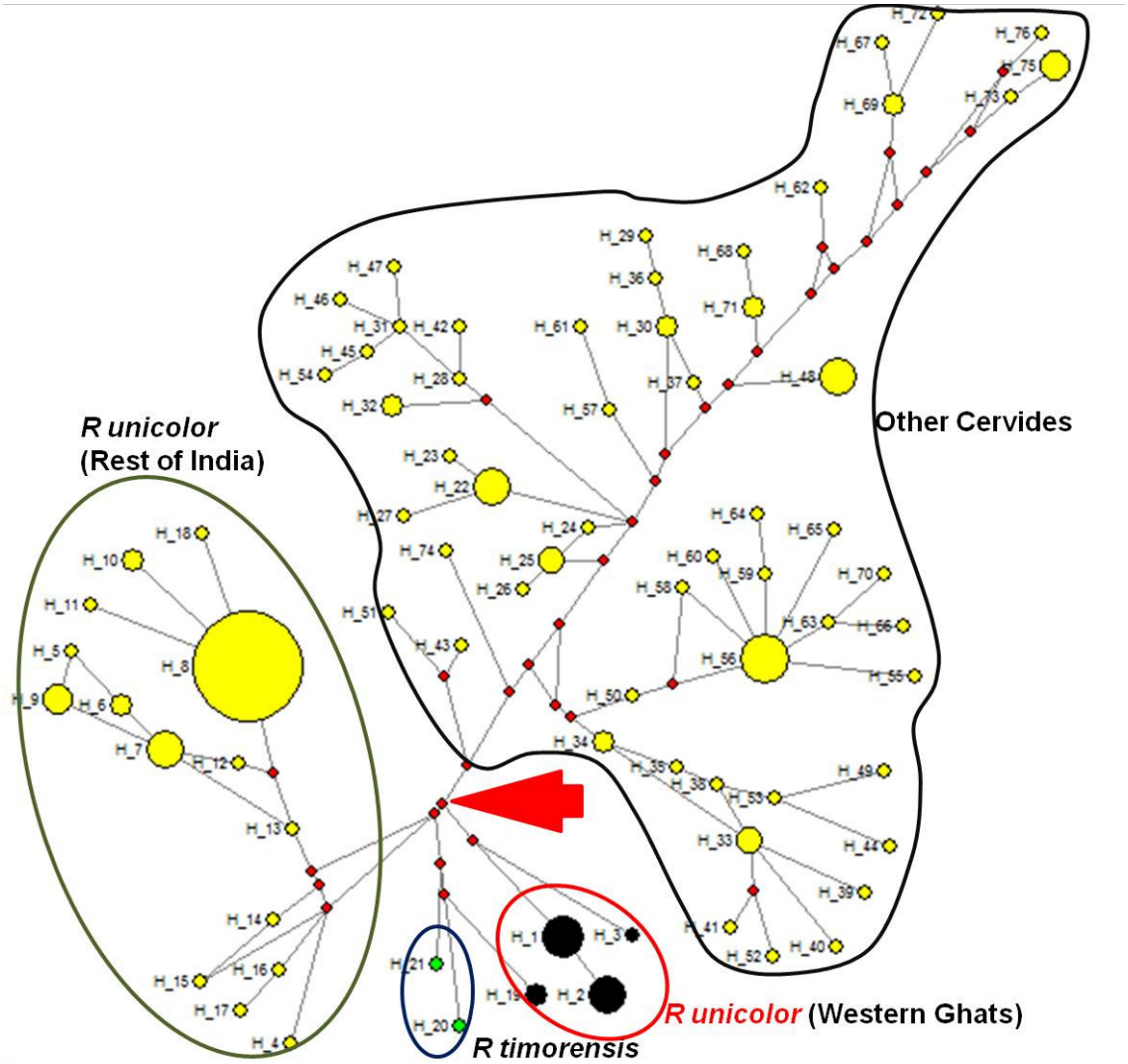


Figure 5.4: Networking of *Rusa spp.* along with the other cervides of world. Arrow indicating the first node of Rusa genus. Yellow hplotypes in green circle represent the main *R unicolor* cluster. Black haplotypes in red circle represent the *R unicolor* of Western Ghats. Green haplotypes in blue circle represent the *R timirensis*. Yellow hplotypes in black circle represent the other cervides spp.

5.4 DISCUSSION

Microsatellite loci developed for several related cervides and bovides (Bishop *et al.* 1994; Vaiman *et al.* 1992; Gaur *et al.* 2003; Poetsch *et al.* 2001; DeWoody *et al.* 1995; Brezinsky *et al.* 1993; Jones *et al.* 2002; Buchanan and Crawford 1993; Marshall *et al.* 1998; Moore *et al.* 1992) were successfully amplified in sambar populations. However, it did not compromise the outcome of the study and a high mean number of alleles per locus (10.74) were observed for examined sambar populations (Table 5.4). It indicated that the markers used were appropriate for assessment of genetic fitness in sambar populations. The mean polymorphic information content (*PIC*) value was also significantly higher (0.711) indicating that tested loci were highly polymorphic for sambar species. The observed heterozygosity was significantly lower than the expected value, which could be an indication of the deviation from expected Hardy-Weinberg saturation.

Multilocus genotyping data indicated that sambar population of Western Ghats have higher genetic diversity with heterozygosity level of 0.533 than those of central India with heterozygosity level of 0.474 (Table 5.2 and 5.3). This may be due to the large connected landscape of Western Ghats. Barriers in landscape connectivity reduce the genetic variation in population (Pérez-Espona *et al.* 2008). However, due to the landscape connectivity issue in Deccan Peninsular biogeographic zone the observed heterozygosity was low.

The populations identified by MCMC based Bayesian analysis suggesting that **(i)** individuals assigned to population 3 (blue bar in K=3 of Figure 5.2) were restricted in Western Ghats of Kerala **(ii)** individuals assigned to reaming two populations (red and green bar in K=3 of Figure 5.2) were observed in Western Ghats and Deccan Peninsula. Structure analysis appeared to be supporting the Gene flow among the populations from north to south rather than south to north. It was already explained that in marked contrast to the Peninsula, the Eastern Amphitheatre shows strong evidence of faunal radiation (Mani 1974). Hence, it could be considered as a genetic support to faunal radiation theory. The data also appeared to be indicating that the restricted population of Western Ghats could probably be peninsular autochthonous elements.

The creation of the Assam gateway provided certainly the most important phase in the biogeographical evolution of India. It facilitated a path for extensive interchange between the Peninsular autochthonous and Asiatic Tertiary-mountain flora and faunas, the movements being equally strong both from the west to east and vice versa (Mani 1974). The tertiary mountains of south China, Indo-China and Thailand and Malaya further south encroached westwards along the Himalaya till the great defile of the river Sutlej and southwards into the Peninsula and Ceylon, which formed a part of the Peninsular mainland of India (Jacob 1949). The Pleistocene glaciations on the Himalaya established another recent ingredient in the biogeographical evolution of India. During this era the temperate Turkmenian elements of the Himalaya extended southwards till the Peninsular south and during the Inter-Glacial times the peninsular elements advanced northwards to the Himalaya. The history of this alternating advance of the Palaearctic and of the Peninsular types is today summarizes in the seasonal oscillations of temperate forms of plants and animals with the warmer Peninsular types in the transitional Peninsular margin that is known as the Indo-Gangetic Plains of India (Mani 1974). These biogeographic elements theoretically supported the consequence of this molecular study and apparently indicating that it helped in radiation of south-east Asian sambar population towards the peninsular region and *vice-versa*.

5.5 CONCLUSIONS

It was apparently indicated from molecular clock estimation that all the *Rusa spp*. were evolved from common ancestor and Western Ghats exclusive population was primitive among them. It was evolved around 2.8 mya which probably lost its connectivity with that south-east Asian population during unfavorable conditions of ice-age (around 2.5 mya). Further upon normalization of climatic conditions at the end of ice-age the south-east Asian counterpart started radiating towards Peninsular India. During this process at the age of 1.2 mya, it diverged in to two species i.e. *R. unicolor* and *R. timorensis*. The modern sambar of Thailand, Malayan and Indian region could probably a result of radiation of this recent clade of south-east Asia. Whereas, the Western Ghats lineage could be an ancient and surviving populations of this group. It was also appeared that the Western Ghats lineage could be an ecologically significant unit (ESU) and may needs attention for proper management of this population.

CHAPTER 6
SYNTHESIS OF FINDINGS AND IMPLICATIONS

SUMMARY

This study was focused on phylogeography and population genetics of sambar (*Rusa unicolor*) in India. Intricate biological samples (faecal pellets, bone, antler and decomposed tissue) were main source of genetic material for this study; hence a protocol was established to extract PCR amplifiable DNA from these samples. Among three evaluated protocols, Gu-HCl based silica binding protocol was uniformly yielded better DNA for further analysis. For amplification of complete cyt *b* and CO*I* gene from DNA extracted from the above intricate samples, two sets of primers were developed. MtDNA control regions sequences were used to examine intra-species variations. After alignment, sequences generated from majority of the Western Ghats samples exhibited a 40 bp insertion. Twenty six distinct haplotypes were identified among all examined populations of India.

Phylogenetic status of sambar was resolved using DNA sequences generated for the mtDNA cyt *b* gene. It segregated the sambar populations of India into two genetically distinct clades. One of the clade was largely distributed in Gengetic Plains, Semi-Arid, Deccan Peninsula and few localities of Western Ghats and the other clade was restricted to Western Ghats. The Western Ghats sambar exhibited approximately 4% variations in cyt *b* gene, which genetically differentiate them from other examined sambar subspecies and rusa deer. It indicated that Western Ghats sambar appeared to be a significant evolutionary unit (SIU). Multilocus genotyping data exhibited that sambar population of Western Ghats have higher genetic diversity with heterozygosity level of 0.533 than those of central India with heterozygosity level of 0.474. The populations identified by MCMC based Bayesian analysis identified three distinct populations among the examined populations of Western Ghats and Deccan Peninsula.

Based on molecular clock estimation using cyt *b* sequence data it was appeared that all the *Rusa spp*. were evolved from common ancestor and Western Ghats exclusive population could be a primitive among them. It was established before the Pleistocene and probably lost its connectivity with that of south-east Asian population during

unfavorable conditions of ice-age glaciations. When climatic conditions become favorable during the end of ice-age the south-east Asian counterpart started radiating towards remaining part of south Asia and speciated around 1.2 mya into *R. unicolor* and *R. timorensis*. The existing sambar populations except the Western Ghats exclusive populations could probably a result of radiation of this recent clade of south-east Asia.

6.1 BACKGROUND

Phylogeography is an advanced method being used in resolving population substructure at geographical scales and fine taxonomic levels for better conservation planning (Manel *et al*. 2003). It depends upon the assessment of genetic variation across large landscape for a species having wide distribution range. Sambar (*R. unicolor*) is one of the most suitable model ungulate for such study. It is a resident of most of the biogeographic zones of India. It was established by genetic analysis that the Javan rusa (*R. timorensis*) of Indonesia is probably the closest living relative of the sambar. However, they do not exhibit reproductive isolation and can naturally interbreed with the sambar and produce fertile hybrids (Idris and Moin 2009). It is still a debatable topic for separation of a subspecies (*philippinus*) of *R. unicolor* as a distinct species *R. marianna* (Grubb 2005; Francis 2009). Hence, accurate phylogenetic resolution was required to know which recognized subspecies or population of sambar has closer relationship with the rusa. These studies highlighted that taxonomy of rusine deer is ambiguous and additional phylogenetic analyses are needed to clarify the phylogeny of rusine deer (Grubb and Groves 1983; Fernández and Vrba 2005). This study was also warranted because genetic analysis indicated that the Javan rusa (*R. timorensis*) of Indonesia is probably the closest living relative of the sambar. It may also be essential to know the existing phylogenetic variations among sambar population of India for identification of population(s) for proper conservation management. In this study, a phylogeography and population genetics was conducted across a large landscape of India for sambar (*R. unicolor*), which has a wide distribution range.

This thesis primarily focused on the examination of intra and inter-species genetic variations among samabr populations in India. During the course of the study a couples of novel methodologies for DNA extraction and amplification were setforth

and new findings pertaining to the genetic variations among sambar population was observed. These finding have been detailed and discussed in Chapter 3, 4 and 5. In this Chapter the synthesis of the methodologies developed and results have been made on the light of similar studies conducted elsewhere.

6.2 DNA EXTRACTION FROM NON-INVASIVE SAMPLES

For successful extraction of amplifiable quality DNA from bones, antlers and faeces of sambar methodologies developed during the course of this study (Gupta *et al.* 2013) was compared with three different DNA extraction protocols viz. Phenol-Chloroform (PC), column based Qiagen kit, and Guanidine hydrochloride (Gu-HCl) based in-house method. The effectiveness of the protocols was compared for higher success rate of PCR amplification from the extracted DNA. The three different DNA extraction protocols yielded different PCR outcomes (Figure 6.1). PCR amplification in bone samples was negligible in PC method, low in Qiagen kit and high with Gu-HCl method. Poor amplification was observed for DNA extracted from bone, antler, and faeces using PC method. In antlers, PCR amplifications were detected in all the three methods. For faeces, Qiagen Stool kit and Gu-HCl method show comparable PCR amplification. The PC method was found inconsistent with the fecal samples. The comparative analysis of the success rate of PCR amplification is shown in Figure 6.2.

This study evidently indicates that two different Qiagen kits (blood/tissue and stool) were required for obtaining the higher PCR success rate from different sample types. However, Gu-HCl based DNA extraction protocol illustrated uniformly higher PCR success from all types of biological samples. The amplified PCR product yields good DNA sequence, which showed authentic matching of the source of its origin and confirmed that the same can be applied on a variety of biological samples with uniform success rate.

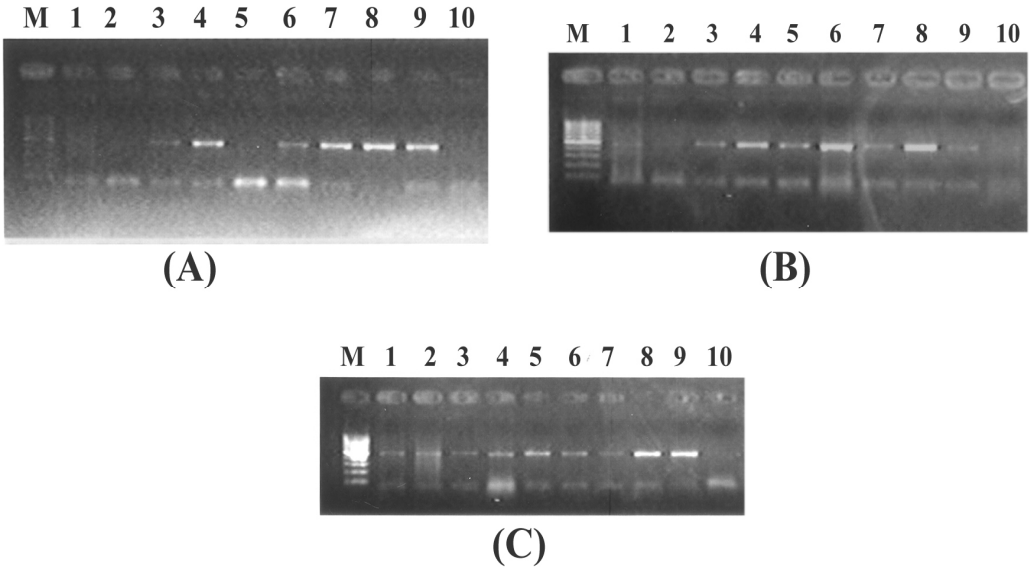


Figure 6.1: PCR amplification with the DNA extracted by (A) Phenol-Chloroform method, (B) Qiagen kit, (C) Gu-HCl and silica binding method. PCR amplification of the DNA extracted from various biological samples as: lane 1- 2 Bone; lane 3-4 Antler; lane 4-5 fecal pallet; lane 6-7 Tissue; lane 9 positive control, and lane 10 negative control

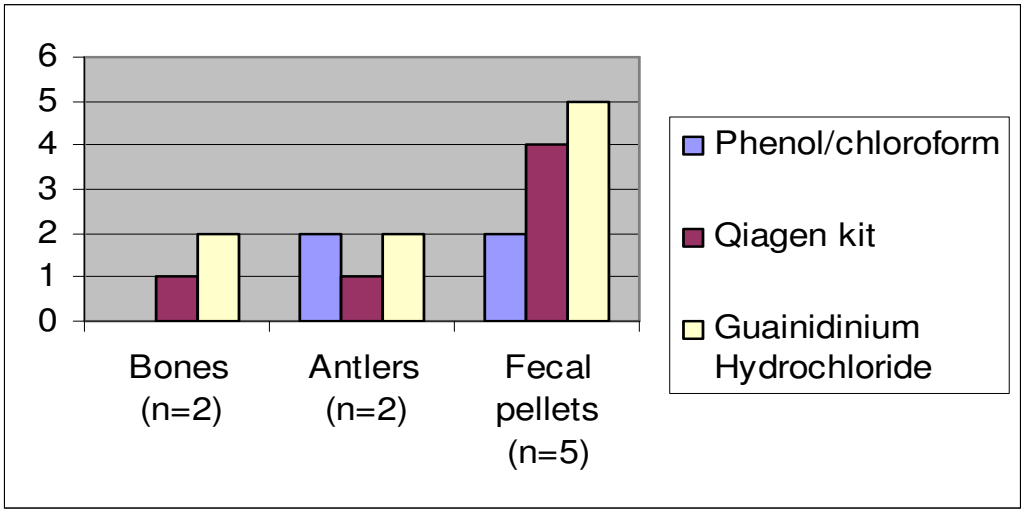


Figure 6.2: Graph showing the PCR amplification index by different extraction methods

This study evidently indicates that two different Qiagen kits (blood/tissue and stool) were required for obtaining the higher PCR success rate from different sample types. However, Gu-HCl based DNA extraction protocol illustrated uniformly higher PCR success from all types of biological samples. The amplified PCR product yields good DNA sequence, which showed authentic matching of the source of its origin and confirmed that the same can be applied on a variety of biological samples with uniform success rate. The uniform success rate in PCR and sequencing evidenced that Gu-HCl method is fast, low-cost, and less hazardous. This protocol was also successfully tested on antler and bone samples of Swamp deer (*Rucervus duvaucelii*) and Eld's deer (*Rucervus eldii eldii*). This study highlights that silica based indigenous DNA extraction protocol using Gu-HCl chaotropic salts yields better quality DNA with higher PCR amplification success rate.

6.3 DEVELOPMENT OF NOVEL PRIMERS FOR CYT *B* GENE

A set of novel primers for successful amplification of the complete mitochondrial cytochrome *b* (cyt *b*) gene of ungulate species was described during this Ph.D. work (Gupta *et al.* 2014). DNA extracted from non-invasively obtained and decomposed samples is found to be degraded and inappropriate for amplification of the complete gene (more than 1 kb) in single PCR amplification. A series of six ungulate-specific conserved primers for the cyt *b* gene were developed for sequencing of complete cyt *b* gene. These primers, in various combinations, amplify 366-1266 bp fragments. This was also used in multiplex PCR reaction to assess the maximum possible size of the PCR amplification product (Table 6.1).

Table 6.1. All possible combinations of the primers during PCR amplification and expected amplicon lengths.

S. No.	Primer combination	Expected Amplicon length(s)
1	MC *b* (14081-105) **F1** and MC *b* (14584-560) **R1**	503 bp
2	MC *b* (14081-105) **F1** and MC *b* (14934-981) **R2**	853 bp
3	MC *b* (14081-105) **F1** and MC *b* (15347-323) **R3**	1266 bp
4	MC *b* (14560-584) **F2** and MC *b* (14934-981) **R2**	374 bp
5	MC *b* (14560-584) **F2** and MC *b* (15347-323) **R3**	787 bp
6	MC *b* (14981-934) **F3** and MC *b* (15347-323) **R3**	366 bp
In multiplex reactions		
7	**F1, R1, R2 and R3**	503, 853 and 1266 bp
8	**F2, R2 and R3**	374 and 887 bp
9	**F1, F2, R2 and R3**	374, 787, 853 and 1266 bp
10	**F1, F2, F3 and R3**	366, 787 and 1266 bp

The mitochondrial genome is a widely preferred locus for the development of conserved primers (Kocher et al. 1989; Verma and Singh 2003). Such primers have proved to be valuable for species identification in the wildlife forensics (Gupta et al. 2005). Recently, a pair of conserved primer for the complete cyt *b* gene was developed for mammals (Naidu et al. 2012). Amplifying the complete cyt *b* gene requires high-quality DNA; however, conservation and forensic geneticists often depend upon non-invasively obtained and forensic (mostly degraded) biological samples. Amplification and analysis of longer fragments (more than 1 kb) from DNA extracted from such samples is challenging. A primer amplifying shorter fragments will be more useful for such studies. A set of novel primers was described for amplification of the complete mtDNA cyt *b* gene of ungulate species from degraded samples. Complete cyt *b* sequences of various ungulate species were obtained from GenBank and aligned using ClustalW (Thompson et al. 1994). Three forward primers and three reverse primers were designed on the basis of sequence homology.

PCR amplicons of expected length were obtained with different primer combinations. Shorter fragments of length 366, 374, and 503 bp could be amplified with the DNA

from all the biological samples, including fecal pellets. A PCR product of length 1266 bp was amplified with the DNA extracted from fresh and less decomposed samples of flesh. However, in the majority of moderately putrefied flesh samples, amplicons of length 853 and 787 bp were obtained. After alignment of the DNA sequences obtained from all the above amplicons, complete gene sequences were obtained for various ungulate species. A complete list of the species tested and the corresponding results are provided in table 6.2.

Table 6.2: List of tested species for validation of cyt *b* primer

S. No.	**Amplification and sequencing with ungulate species**	**No. of individuals**
1	Sambar (*Rusa unicolor*)	9
2	Swamp deer (*Rucervus duvaucelii*)	2
3	Chital (*Axis axis*)	3
4	Hog deer (*Axis porcinus*)	3
5	Barking deer (*Muntiacus muntjak*)	2
6	Black buck (*Antelope cervicapra*)	2
7	Indian gazelle (*Gazella bennettii*)	2
8	Nilgai (*Boselaphus tragocamelus*)	4
9	Indian wild pig (*Sus scrofa*)	8
10	Domestic pig (*Sus scrofa*)	6

6.4 DEVELOPMENT OF NOVEL PRIMERS FOR CO*I* GENE

Similar to above, a set of novel primers for successful amplification of the complete mitochondrial cytochrome oxidase *I* (CO*I*) gene of ungulate species was also described during this dissertation work. A series of six ungulate-specific conserved primers for the cyt *b* gene were developed for sequencing of complete CO*I* gene. These primers, in various combinations, amplify 301-1599 bp fragments. I also used them in multiplex PCR reaction to assess the maximum possible size of the PCR amplification product.

Amplifying the complete CO*I* gene requires high-quality DNA, which is a challenging task in conservation and forensic geneticists (DNA from degraded biological

samples). A primer amplifying shorter fragments will be more useful for such studies. I describe novel primers for amplification of the complete mtDNA cyt *b* gene of these animals from degraded samples. Complete CO*I* sequences of various ungulate species were obtained from GenBank and aligned using ClustalW (Thompson et al. 1994). Three forward primers and three reverse primers were designed on the basis of sequence homology.

Expected PCR amplicons were obtained with different primer combinations. Less than 1 KB fragments (301, 475, 754 and 837 bp) could be amplified with the DNA from all the biological samples, including hair samples. PCR products of length 1320 and 1599 bp were amplified with the DNA extracted from fresh and less decomposed flesh. After alignment of DNA sequences obtained from all the above amplicons, a complete gene consensus sequences were obtained for various ungulate species table 6.2.

Table 6.3: List of tested species for validation of CO*I* primer

S. No.	Amplification and sequencing with ungulate species	No. of individuals
1	Sambar (*Rusa unicolor*)	5
2	Swamp deer (*Rucervus duvaucelii*)	2
3	Chital (*Axis axis*)	2
4	Hog deer (*Axis porcinus*)	2
5	Barking deer (*Muntiacus muntjak*)	2
6	Black buck (*Antelope cervicapra*)	2
7	Indian gazelle (*Gazella bennettii*)	2
8	Nilgai (*Boselaphus tragocamelus*)	2
9	Indian wild pig (*Sus scrofa*)	5
10	Domestic pig (*Sus scrofa*)	4

6.5 EXAMINATION OF INTRA-SPECIES VARIATION

Despite its large distribution and significant morphological variation no study has been conducted to examine the genetic variation among sambar (*Rusa unicolor*) populations in India. This dissertation describes genetic variation among select sambar populations using Mitochondrial DNA (mtDNA) control region. Sequence variation in

a partial fragment of mtDNA control region was examined from the biological samples collected from Southern India, Central highlands and Semi-Arid regions. This study demonstrated that the most of the South Indian population is distinct from entire central highland, north and several south India populations in a distinct genetic feature.

PCR amplifications that were approximately 40 bp longer than expected were obtained with DNA extracted from 23 samples from south India (Fig. 3.2). Three samples from south India and all 38 samples from central and north India yielded amplifications of the expected size. This indicates that the majority of the animals of the south Indian population have a unique feature in the mtDNA control region. Using the primer pair (Balakrishnan et al. 2003), 543–623 bp sequences of the control region were obtained. These were aligned for further analysis. After alignment, the sequences generated from the 23 samples of Western Ghats exhibited a 40 bp insertion after nucleotide (nt) position 233 (Table 3.1). This unique insertion in a majority of the Western Ghats samples indicates that there is a significant genetic variation in this population. Four out of the 23 insertion-positive Western Ghats samples were collected from a confined population from a zoological park in Goa. These individuals had the sore spot on the ventral region of the neck. The 40 bp insertion was absent in three samples of south India, where some populations do not possess the sore patch. The 40 bp insertion was not found in the samples from the animals 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 of central and north Indian populations. Overall, the mtDNA control region demonstrates high variability among sambar populations, and 26 distinct haplotypes were identified (Table 3.1). The haplotype (RUC1–RUC26) sequences were submitted to GenBank (accession numbers KF133981-99 and KF648589-95, Table 3.1). Haplotypes RUC1–RUC5, RUC8–RUC15, RUC20 and RUC21 are 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 central India and RUC19 in three individuals from central India (Table 3.1). The pairwise distance and percentage similarity matrices showed wide ranges of 0.002–0.113 and 91.60–99.83, respectively (Table 3.2).

6.6 MOLECULAR PHYLOGENY OF SAMBAR

Phylogenetic status of sambar (*R. unicolor*) population in India was described in this thesis. The DNA sequences were generated for the mtDNA cyt *b* gene from sambar population across India. Analysis of phylogenetic status based on cyt *b* gene sequence demonstrated the unambiguous genetic variation among sambar populations of India and segregated them into two genetically distinct clades. One of the clade was largely distributed in Gengetic Plains, Semi-Arid, Deccan Peninsula and few localities of Western Ghats and the other clade was restricted to Western Ghats.

Phylogenetic analysis indicated that there was a wide distribution of one common ancestor of sambar around 2.8 million year ago (mya). This ancient ancestor was fragmented and restricted to Western Ghats in India and Javan islands. The modern sambar (*Rusa unicolor*) evolved from the Javan sambar and regained its ancestor's habitat in south and south-east Asia. The sambar population reached to north-east India and further moves across the India and gained genetic changes. The Western Ghats sambar did not go to evolutionary changes and retained its habitat. Hence, India inhabited two distinct lineage of sambar one is similar to south-east Asian lineage and other is genetically distinct Western Ghats lineage. Moreover, the Western Ghats sambar exhibit around 5% variations in cyt *b* gene, which genetically differentiate them from other sambar subspecies. Hence Western Ghats sambar appeared to be significant evolutionary unit.

6.7 POPULATION GENETICS AND GENE FLOW

Multilocus genotyping data indicating that sambar population of Western Ghats have higher genetic diversity with heterozygosity level of 0. 0.533 than those of central India with heterozygosity level of 0.474. This may be due to the large connected landscape of Western Ghats. However, due to connectivity issue in Deccan Peninsula landscape the observed heterozygosity was low. The populations identified by MCMC based Bayesian analysis suggesting that (i) individuals assigned to population 3 (blue bar in K=3 of Figure 5.2) were restricted in Western Ghats of Kerala (ii) individuals assigned to reaming two populations (red and green bar in K=3 of Figure 5.2) were observed in Western Ghats and Deccan Peninsula.

It was apparently indicated from molecular clock estimation based on cyt *b* sequence data that all the *Rusa spp*. were evolved from common ancestor and Western Ghats exclusive population could be a primitive among them. It was evolved around 2.8 mya which probably lost its connectivity with that south-east Asian population during unfavourable conditions of ice-age (around 2.5 mya). Further upon normalization of climatic conditions at the end of ice-age the south-east Asian counterpart started radiating and speciation in to *R. unicolor* and *R timorensis* around 1.2 mya. The modern sambar of Thailand, Malayan and Indian region could probably a result of radiation of this recent clade of south-east Asia. Whereas, the Western Ghats clade could probably be ancient and surviving populations of this group.

6.8 CONCLUSIONS

Other than the main three objectives set forth for the assessment of genetic variations among sambar populations, this study also described the comparison of three DNA extraction protocols for successful extraction of PCR amplifiable quality DNA from bones, antlers and faeces samples of sambar. Three different DNA extraction protocols were compared in this study including Phenol-Chloroform (PC), column based Qiagen kit, and Guanidine hydrochloride (Gu-HCl) based in-house method. The effectiveness of the protocols was compared for higher success rate of PCR amplification from the extracted DNA. It was highlights that silica based indigenous DNA extraction protocol using Gu-HCl chaotropic salts yields better quality DNA with higher PCR amplification success rate.

DNA extracted from non-invasively obtained and decomposed samples is found to be degraded and inappropriate for amplification of the complete gene (more than 1 kb) in single PCR amplification. A set of novel primers were also established for successful amplification of the complete mitochondrial cytochrome *b* (cyt *b*) gene of ungulate species. During this study, a series of six ungulate-specific conserved primers were developed for the amplification and sequencing of complete cyt *b* gene. These primers, in various combinations, amplify 366-1266 bp fragments. These primers were also used in multiplex PCR reaction to obtain the maximum possible size of the PCR amplification product.

6.8.1 Intra-species genetic variations

In this study, control region was used to evaluate the genetic variations among sambar population in India. An interesting Insertion–deletion (INDEL) was observed among studied sambar population, which were commonly used in human genetics and forensics (Thangaraj *et al*. 2006; da Costa Francez *et al*. 2012). The detailed outcome was discussed in chapter 3 section 3.4. INDEL of 40 bp long stretch was detected in the sambar populations of geographically distinct locations across India (Figure 3.2). This unique 40 bp insertion was detected in select population of Western Ghats, which was absent in the sambar populations of Gangetic Plains, Semi-Arid, Deccan Peninsula and few individuals of area located at the fringes of Western Ghats and Deccan Peninsula. This study also scientifically highlighted the existence of one ecologically important molecular signature (40 bp insertion at nt position 233) in select sambar population. Two genetically distinct populations of sambar with excess of haplotype variations (total 26 haplotype) were identified for proper wildlife management.

6.8.2 Molecular taxonomy of sambar

It indicated that the sambar population in Western Ghats is ancient lineage, which probably did not participated in active evolution process and refrain wide distribution so as to restrict themselves in the vicinity of Western Ghats. The Javan ancestor further evolved into two distinct species of genus Rusa i.e. *R. unicolor* and *R. timorensis.* It was assumed that south-east Asia could be a centre for speciation of modern *R. unicolor* (around 1 mya). More sampling from this region is required to reach a firm conclusion. However, this study appeared to be indicating that the Western Ghats lineage is genetically distinct from remaining *R. unicolor* and *R. timorensis.* The select Western Ghats sambar appeared to be ancient among extant rusa and could probably be an evolutionary significant unit, which has significant older divergence time of 2.8 mya than the remaining *R. unicolor* and *R. timorensis* (around 1 mya).

6.8.3 Population genetics of sambar

Multilocus genotyping data indicated that sambar population of Western Ghats have higher genetic diversity with heterozygosity level of 0.533 than those of central India with heterozygosity level of 0.474 (Table 5.2 and 5.3). This may be due to the

large connected landscape of Western Ghats. However, due to the landscape connectivity issue in Deccan Peninsular biogeographic zone the observed heterozygosity was low. The populations identified by MCMC based Bayesian analysis suggesting that (**i)** individuals assigned to population 3 (blue bar in K=3 of Figure 5.2) were restricted in Western Ghats of Kerala (**ii)** individuals assigned to reaming two populations (red and green bar in K=3 of Figure 5.2) were observed in Western Ghats and Deccan Peninsula. Structure analysis appeared to be supporting the Gene flow among the populations from north to south rather than south to north. It was already explained that in marked contrast to the Peninsula, the Eastern Amphitheatre shows strong evidence of faunal radiation (Mani 1974). Hence, it could be considered as a genetic support to faunal radiation theory. The data also appeared to be indicating that the restricted population of Western Ghats could probably be peninsular autochthonous elements.

It was apparently indicated from molecular clock estimation that all the *Rusa spp*. were evolved from common ancestor and Western Ghats exclusive population was primitive among them. It was evolved around 2.8 mya which probably lost its connectivity with that south-east Asian population during unfavourable conditions of ice-age (around 2.5 mya). Further upon normalization of climatic conditions at the end of ice-age the south-east Asian counterpart started radiating towards Peninsular India. During this process at the age of 1.2 mya, it diverged in to two species i.e. *R. unicolor* and *R. timorensis*. The modern sambar of Thailand, Malayan and Indian region could probably a result of radiation of this recent clade of south-east Asia. Whereas, the Western Ghats lineage could be an ancient and surviving populations of this group. It was also appeared that the Western Ghats lineage could be an ecologically significant unit (ESU) and may need further investigation and conservation action.

REFERENCES

Acharjyo LN, Misra R (1971) Age at sexual maturity of three species of wild animals in captivity. Journal of the Bombay Natural History Society **68**:446.

Alexander SM, Waters NM (2000) The effects of highway transportation corridors on wildlife: a case study of Banff National Park. Transportation Research **8:**307–320.

Allen GM (1940) The mammals of China and Mongolia. Natural history of central Asia. Vol. XI. Part 2. American Museum of Natural History, New York.

Allendorf FW, Luikart G (2007) Conservation and the genetics of populations. Blackwell Publishing, Malden, Massachusetts.

Alvarez Y, Mateo JA, Andreu AC, Diaz–Paniagua C, Diez A, Bautista JM (2000) Mitochondrial DNA haplotyping of Testudo graeca on both continental sides of the Straits of Gibraltar. Journal of Heredity **91**:39–41.

Anh PT, Khoi ND, Long NTT (1996) Checklist of mammal collection of Da Lat Biological Subinstitute. Tap Chi Sinh Hoc (J. Biology, Hanoi) **18**: 16-22.

Antin JH, Childs R, Filipovich AH Giralt S, Mackinnon S, Spitzer T, Weisdorf D (2001) Establishment of complete and mixed donor chimerism after allogeneic lymphohematopoietic transplantation: recommendations from a workshop at the 2001 Tandem Meetings of the International Bone Marrow Transplant Registry and the American Society of Blood and Marrow Transplantation. Biology of Blood and Marrow Transplantation **7**:473–85.

Antonelli A, Sanmartin I (2011) Why are there so many plant species in the Neotropics? Taxon **60**:403–414.

Aragon E, Goin FJ, Aguilera YE, Woodburne MO, Carlini AA, Roggiero MF (2011) Palaeogeography and palaeoenvironments of northern Patagonia from the Late Cretaceous to the Miocene: the Palaeogene Andean gap and the rise of the North Patagonian High Plateau. Biological Journal of the Linnean Society **103**:305–315.

Arif IA, Khan HA (2009) Molecular markers for biodiversity analysis of wildlife animals: a brief review. Animal Biodiversity and Conservation **32**:9-17.

Asher GW, Muir PD, Semiadi G, O'neill KT, Scott IC, Barry TN (1997) Seasonal patterns of luteal cyclicity in young red deer (*Cervus elaphus*) and sambar deer (*Cervus unicolor*). Reproduction, Fertility and Development **9**:587–596.

Avise JC (1994) Molecular markers, natural history and evolution Chapman and Hall, New York.

Avise JC (1998) The history and purview of phylogeography: a personal reflection. Molecular Ecology **7**:371–379.

Avise JC (2000) Phylogeography: The History and Formation of Species. President and Fellows of Harvard College.

Avise JC (2004) Molecular markers, natural history and evolution. Sinauer Associates, Sunderland, Massachusetts.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annual Review of Ecology and Systematics **18**:489-522.

Bacon A-M, *et al.* (2004) The Pleistocene Ma U'Oi cave, northern Vietnam: palaeontology, sedimentology, and palaeoenvironments. Geobios **37:**305–314.

Bagchi S, Goyal SP, Sankar K (2003) Prey abundance and prey selection by tigers (*Panthera tigris*) in a semi-arid, dry deciduous forest in western India. Journal Zoology **260**:285-290.

Bagchi S, Goyal SP, Sankar K (2004) Herbivore density and biomass in a semi-arid tropical dry deciduous forest of western India. Journal of Tropical Ecology **20**:475-478.

Baker SW (1898) Wild beasts and their ways: reminiscences of Europe, Asia, Africa, and America. Macmillan and Company, Limited, New York.

Balakrishnan CN, Monfort SL, Gaur A *et al.* (2003) Phylogeography and Conservation Genetics of Eld's deer (*Cervus eldi*). Molecular Ecology **12:**1-10.

Barrett SCH, Kohn J (1991) The genetic and evolutionary consequences of small population size in plant: implications for conservation In: Genetics and Conservation of Rare Plants (Eds D Falk and KE Holsinger), pp 3-30 Oxford University Press.

Barton N, Clark A (1990) Population structure and process in evolution in K Wohrmann and S K Jain (Eds) Population biology: ecological and evolutionary viewpoint *Springer*, Berlin, pages 115-173.

Basbar MA, Begum F, Mondal KK (2001) Sambar deer breeding in captivity in Dulahazra Safary (Managed National Reserve), Bangladesh. Tigerpaper **28**:21–23.

Behling H (2002) South and southeast Brazilian grasslandsduring Late Quaternary times: a synthesis. Palaeogeography Palaeoclimatology Palaeoecology **177**:19–27.

Bentley A (1978) An introduction to the deer of Australia Koetong ed Koetong Trust Service, Victoria, Australia

Berwick SH (1974) The community of wild ruminants in the Gir Forest ecosystem, India. Ph.D. dissertation, Yale University, New Haven, Connecticut.

Bishop MD, Kappe SM, Keele JW, Stone RT, Sunden SL, Hawkins GA, Toldo SS, Fries R, Grosz MD, Yoo J (1994) A linkage map for cattle. Genetics **136**:619–639.

Biswas S, Sankar K (2002) Prey abundance and food habit of tigers in Pench National Park, Madhya Pradesh, India. Journal of Zoology **256**:411-420

Blanford,WT (1888) The fauna of British India, including Ceylon and Burma. Taylor and Francis, London, United Kingdom.

Blouin MS, Parsons M, Lacaille V, Lotz, S (1996) Use of microsatellite loci to classify individuals by relatedness. Molecular Ecology **5**:393–401.

Boitani L (1984) Simon & Schuster's Guide to Mammals Simon & Schuster/Touchstone Books.

Bonnet X, Shine R, Lourdais O (2002) Taxonomic chauvinism. Trends in Ecology and Evolution **17:**1–3.

Bradshaw Foundation Mitochondrial DNA: The Eve Gene. Bradshaw Foundation. Retrieved 5 November 2012.

Brander AAD (1923) Wild animals in central India. Edward Arnold & Company, London, United Kingdom.

Braun A, Groves CP, Grubb P, Yang Q-S, Xia L (2001) Catalogue of the Musée Heude collection of mammal skulls. Acta Zootaxonomica Sinica **26**:608–660.

Brezinsky L, Kemp SJ, Teale AJ (1993) ILSTS005: a polymorphic bovinemicrosatellite. Animal Genetics **24**:75-76.

Brown JH, Lomolino MV (1998) Biogeography (2nd edn), Sinauer Associates.

Brown JR, Beckenbach AT, Smith MJ (1993) Interspecific DNA sequence variation of the mitochondrial Control Region of White Sturgeon. Molecular Biological Evolution **10**:326-341.

Buchanan EC, Crawford AM (1992) Ovine dinucleotide repeat polymorphism at the MAF70 locus. Animal Genetics **23**:185.

Burnie D, Wilson DE (2005)) Animal: The Definitive Visual Guide to the World's Wildlife DK Adult.

Butler JM (2005) Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers, Second Edition, Academic Press.

Cardillo M, Mace GM, Gittleman JL, Purvis A (2006) Latent extinction risk and the future battle grounds of mammal conservation. Proceedings of the Natural Academy of Sciences, USA **103**:4157–4161.

Carson HL, Templeton AR (1984) Genetic revolutions in relation to speciation phenomena: The founding of new populations. Annu Rev Ecol Syst **15:**97–131.

Castella V, Ruedi M, Excoffier L Ibañez C, Arlettaz R, Hausser J (2000) Is the Gibraltar Strait a barrier to gene flow for the bat Myotis myotis (Chiroptera: Vespertilionidae)? Molecular Ecology **9:**1761–1772.

Ceballos G, Ehrlich PR, Soberon J, Salazar I, Fay JP (2005) Global mammal conservation: what must we manage? Science **309**:603–607.

Chan JP-W, Tsai H-Y, Chen C-F, Tung K-C, Chang C-C (2009) The reproductive performance of female Formosan sambar deer (*Cervus unicolor swinhoei*) in semi-domesticated herds. Theriogenology **71**:1156–1161.

Channell R, Lomolino M (2000) Dynamic biogeography and the conservation of endangered species. Nature **403**:84–86.

Chatterjee H, Ho S, Barnes I, Groves C (2009) Estimating the phylogeny and divergence times of primates using a supermatrix approach. BMC Evolutionary Biology **9**:259.

Cipriani G, Lot G, Huang WG, Matarazzo MT, Peterlunger E and Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: Isolation, characterisation and cross species amplification in*Prunus*. Theoretical and Applied Genetics **99**:65-72.

Clement M, Posada D, Crandall KA (2000) TCS: A Computer Program to estimate gene genealogies. Molecular Ecology **9**:1657 – 1659.

Clutton-Brock TH, Albon SD (1989) Red Deer in the Highlands. BSP Professional Books, Oxford, UK.

Clutton-Brock TH, Albon SD, Harvey PH (1980) Antlers, body size, and breeding group size in the Cervidae. Nature **285**:565–567.

Colinvaux PA, De Oliveira PE, Moreno JE, Miller MC, Bush MB (1996) A long pollen record from lowland Amazonia: forest and cooling in glacial times. Science **274**:85–88.

Collevatti RG, Brondani RV and Grattapaglia D (1999) Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. Heredity **83**:748-756.

Collevatti RG, Grattapaglia D, Hay JD (2001) Population genetic structure of the endangered tropical tree species Caryocar brasiliense, based on variability at microsatellite loci. Molecular Ecology **10:**349-356,.

Comber E (1904) The breeding season of big game. Journal of the Bombay Natural History Society **16**:176–179.

Compagnucci HR (2011) Atmospheric circulation over Patagonia from the Jurassic to present: a review through proxy data and climatic modelling scenarios. Biological Journal of the Linnean Society **103**:229–249.

Conry PJ (1988) Management of feral and exotic game species on Guam. Transactions of the Western Section of The Wildlife Society **24**:26–30.

Constable JL, Ashley MV, Goodall J, Pusey AE (2001) Non-invasive paternity assignment in Gombe chimpanzees. Molecular Ecology **10**:1279-1300.

Corbet GB, Hill JE (1992) The mammals of the Indomalayan region: a systematic review. Oxford University Press, Oxford, United Kingdom.

Coulon A, Cosson J-F, Morellet N, Angibault J-M, Cargnelutti B, Galan M, Aulagnier S, Hewison AJM (2006a) Dispersal is not female biased in a resource-defence mating ungulate, the European roe deer. Proceedings of the Royal Society of London, Series B **273:**341-348.

Coulon A, Guillot G, Cosson J-F, Angibault JM, Aulagnier S, Cargnelutti B, Galan M, Hewison AJM (2006b) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. Molecular Ecology **15**:1669-1679.

Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison AJM (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape : an individual-based approach. Molecular Ecology **13**:2841-2850.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution **17:**390–395.

Crandall L (1964) The management of wild animals in captivity University of Chicago press, Chicago.

Crnokrak P, Roff DA (1999) Inbreeding depression in the wild. Heredity **83**:260–270.

Cruzan MB, Templeton AR (2000) Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. Trends in Ecology and Evolution **15**:491–496.

Cummings CJ and Zoghbi HY (2000) Trinucleotide repeats: Mechanisms and pathophysiology. Annual Review of Genomics and Human Genetics **1**:281-328.

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. American Naturalist **168:**486–489.

da Costa Francez PA, Rodrigues EM, de Velasco AM, dos Santos SE (2012) Insertion-deletion polymorphisms--utilization on forensic analysis. International Journal of Legal Medicine **126**:491-496.

Daniels M, McClean C (2003) Red deer calf tagging programmes in Scotland — an analysis. Deer (The Journal of the British Deer Society) **12**:420–423.

Danny P (2002) Mitochondria can be inherited from both parents. NewScientist.com. Retrieved 2008-02-05.

Davarn S (1938) Cause of sore neck in sambar. Journal of the Bombay Natural History Society **40**:118–122.

Davidson WR, Blue JL, Flynn LB, Shea SM, Marchinton RL, Lewis JA (1987) Parasites, diseases and health status of sympatric populations of sambar deer and white-tailed deer in Florida. Journal of Wildlife Diseases **23**:267–272.

Davidson, M.N., 1983. Sika Deer. In: The Handbook of New Zealand Mammals, King, C.M., (Ed.). Oxford University Press, USA. 468-477.

Dayanandan S, Bawa KS and Kesseli R (1997) Conservation of microsatellites among tropical tree (Leguminosae). American Journal of Botany **84**:1658-1663.

De Queiroz A (2005) The resurrection of oceanic dispersal in historical biogeography. Trends in Ecology and Evolution **20**:68–73.

DeWoody JA, Honeycutt RL, Skow LC (1995) Microssatellite markers in white-tailed deer. Journal of Heredity **86**:317-319.

Di Stefano G, Petronio C (2002) Systematics and evolution of the Eurasian Plio-Pleistocene tribe Cervini (Artiodactyla, Mammalia). Geologica Romana **36**:311–334.

Di Stefano G, Petronio C (2002) Systematics and evolution of the Eurasian Plio-Pleistocene tribe Cervini (Artiodactyla, Mammalia). *Geologica Romana* **36:**311–334.

Dinerstein E (1979) An ecological survey of the Royal Karnali-Bardia Wildlife Reserve, Nepal Part II:Habitat/animal interactions. Biological Conservation **16:**265-300.

Dong W (1993) The fossil record of deer in China. In: Deer in China: biology and management (Ohtaishi N, Sheng HL eds.). Elsevier Science Publishers, Amsterdam, The Netherlands. Pp. 95–102

Douzery E, Randi E (1997) The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic content. Molecular Biology and Evolution **14**:1154-1166.

Downes, M (1983) The forest deer project 1982 Australian Deer Research Foundation Ltd, Melbourne, Australia

Driscoll CA, Nelson G, Goldstein D, O' Brien SJ (2002) Genomic microsatellites as evolutionary chronometer: A test in Wild Cats. Genome **12**:414-423.

Drummond AJ, Suchard MA, Dong Xie, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution **29:**1969-1973.

Duckworth J, Salter WRE, Khounboline K (compilers) (1999) Wildlife in Lao PDR: 1999 Status Report. Vientiane: IUCN - The World Conservation Union / Wildlife Conservation Society/Centre for Protected Areas and Watershed Management.

Durrett R (2008) Probability Models for DNA Sequence Evolution. Springer

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources **4**:359-361.

Eisenberg J F (1987) The evolutionary history of the Cervidae with special reference to the South American radiation. In: Biology and management of the Cervidae (Wemmer CM, ed.). Smithsonian Institution Press, Washington, D.C. Pp. 60–64

Eisenburg JF, Lockhart M (1972) An ecological reconnaissance of Wilpattu National Park, Ceylon. Smithsonian Contribution to Zoology **101**:1–118.

El-Jaafari HAA, Jothi MP, Ismail I, Siti SS (2008) RAPD Analysis of Three Deer Species in Malaysia Asian-Australian Journal of Animal Science **21**:1233-1237.

Epps CW, Palsboll PJ, Wehausen JD, Roderick GK, Ramey RR, McCullough DR (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. Ecology Letters **8**:1029–1038.

*Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology **14:**2611–2620.*

Evans GP (1912) Big-game shooting in Upper Burma. Longmans, Green, and Company, London, United Kingdom.

Everett CM, Wood NW (2004) Trinucleotide repeats and neurodegenerative disease. Brain **127**:2385-2405.

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annual Review of Ecology and Evolution **34**:487–515

Fernandez HM, Vrba ES (2005) A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. Biological Review **80**:269–302.

Fernando PT, Vidya TNC, Pajapakse C, Dangolla A, Melnick DJ (2003) Reliable non-invasive genotyping: fantasy or reality. Journal of Heredity **94**:115–123.

Field D, Wills C (1998) Abundant microsatellite polymorphism in Saccharomyces cerevisiae, and the different distributions of microsatellites in eight prokaryotes and S. cerevisiae, result from strong mutation pressures and a variety of selective forces. Proceedings of the National Academy of Sciences USA **95**:1647–1652.

Flerov KK (1952) Musk deer and deer. Fauna of USSR. Mammals. Vol. 1. No. 2. National Science Foundation and Smithsonian Institution, Washington, D.C.

Francis CM (2009) A guide to the mammals of southeast Asia. Princeton University Press, Princeton, New Jersey.

Frankham R (1995a) Effective population size/adult population size ratios in wildlife: a review. Genetical Research **66**:95-107

Frankham R (1995b) Inbreeding and extinction: a threshold effect. Conservation Biology **9**:792-799.

Frankham R (2003) Genetics and conservation boilogy. Comptes rendus Biologies, **326**:S22–S29.

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation Genetics. Cambridge Uni- versity Press.

Gaur A, Singh A, Arunabala V, Umapathy G, Shailaja K, Singh L (2003) Development and characterization of 10 novel microsatellite markersfrom chital deer (*Cervus axis*) and their cross-amplification in otherrelated species. Molecular Ecology Notes **3**:607–609.

Geist V (1998) Deer of the world: their evolution, behaviour, and ecologyStackpole Books 1998 pp 73–77.

Gerber AS, Loggins R, Kumar S, Dowling TE (2001) Does nonneutral evolution shape observed patterns of DNA variation in animal mitochondrial genomes? Annual Reviews of Genetics **35**:539–566.

Geyer CJ, Ryder OA, Chemnick LG, Thompson EA (1993) Analysis of relatedness in the California Condor, from DNA fingerprints. Molecular Biology and Evolution **10**:571–589.

Gilbert C, Ropiquet A, Hassanin A (2006) Mitochondrial and nuclear phylogenies of Cervidae (Mammalia, Ruminantia): systematics, morphology, and biogeography. Molecular Phylogenetics and Evolution **40:**101–117.

Gilbert R (1888) Notes on sambhur and sambhur stalking. Journal of the Bombay Natural History Society **3**:224–232.

Glaubitz JC (2004) CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. Molecular Ecology Notes **4**:309-310.

Goldstein DB, Schlotterer C (1999) Microsatellites: Evolution and Applications. Oxford University Press, New York, Pp 343.

Gonzalez S, Maldonado JE, Leonard JA et al. (1998) Conservation genetics of the endangered Pampas deer (*Ozotoceros bezoarticus*). Molecular Ecology **7**:47–56.

González-Ittig RE, Rossi-Fraire HJ, Cantoni GE, Herrero ER, Benedetti R, Gallardo MH, Gardenal CN (2010) Population genetic structure of long-tailed pygmy rice rats (Oligoryzomys longicaudatus) from Argentina and Chile based on the mitochondrial control region. Can. J. Zool. **88:**23–35.

Goodwin BJ (2003) Is landscape connectivity a dependent or independent variable? Landscape Ecology **18**:687–699.

Green MJB (1985) Aspects of the ecology of the Himalayan musk deer. Ph.D. dissertation, University of Cambridge, Cambridge, United Kingdom.

Gribbin JR (1982) Future Weather: Carbon Dioxide, Climate and the Greenhouse Effect. Penguin.

Griffiths AJF, Miller JF, Suzuki,DT, Lewontin RC, Gelbart WM (1996) Introduction to Genetic Analysis, 5th Edition. W.H. Freeman, New York.

Groves CP (2003) Taxonomy of ungulates of the Indian subcontinent.Journal of the Bombay Natural History Society **100**:314–362.

Groves CP (2007) Family Cervidae. In: The evolution of artiodactyls (D. R. Prothero and S. E. Foss, eds.). Johns Hopkins University Press, Baltimore, Maryland. Pp. 249–256

Groves CP, Grubb P (1987) Relationships of living deer. In Biology and management of the Cervidae: 21–59. Wemmer, C. M. (Ed.). Washington, DC: Smithsonian Institution Press.

Grubb P (1990) Cervidae of Southeast Asia. Pp. 169–179 in Horns, pronghorns, and antlers (G. A. Bubenik and A. B. Bubenik, eds.). Spring Verlag, New York.

Grubb P (2005) Order Artiodactyla. in Mammal species of the world: a taxonomic and geographic reference (Wilson DE and Reeder DM ed). 3rd ed. Johns Hopkins University Press, Baltimore, Maryland.Guha S, Goyal SP, Kashyap VK (2006) Genomic variation in the mitochondrially encoded cytochrome *b* (MT–CYB) and 16S rRNA (MT–RNR2) genes: characterization of eight endangered Pecoran species. Animal Genetics **37**:262–265.

Gupta SK, Kumar A and Hussain SA (2014) Novel primers for sequencing of the complete mitochondrial cytochrome b gene of ungulates using non-invasive and degraded biological samples. Conservation Genetics Resources. *In press* (DOI: 10.1007/s12686-014-0143-6).

Gupta SK, Kumar A, Hussain SA (2013) Extraction of PCR-amplifiable DNA from a variety of biological samples with uniform success rate. Conservation Genetics Resources **5**:215-217.

Gupta SK, Kumar A, Hussain SA, Vipin and Singh L (2013b) Cytochrome b based genetic differentiation of Indian wild pig (*Sus scrofa cristatus*) and domestic pig (*Sus scrofa domestica*) and its use in wildlife forensics. Science and Justice **53**:220-222.

Gupta SK, Thangaraj K, Singh L (2011) Identification of the source of ivory idol by DNA analysis. Journal of Forensic Science **56**:1343-1345.

Gupta SK, Verma SK, Singh L (2005) Molecular insight into a wildlife crime: the case of a peafowl slaughter. Forensic Science International **154:**214-217.

Gupta SK, Verma SK, Singh L (2006) Molecular insight into a wildlife crime: a case of peafowl slaughter. Forensic Science International **154**:214-217.

Gururaja KV Aravind NA, Ali S, Ramachandra TV, Velavan TP, Krishnakumar V, Aggarwal RK (2007) A new frog species from the central Western Ghats of India, and its phylogenetic position. Zoological Science **24:**525–534.

Gyllensten U, Wharton D, Josefsson A, Wilson AC (1991). Paternal inheritance of mitochondrial DNA in mice. Nature **352**:255–7.

Haig SM, Ballou JD, Casna NJ (1994) Identification of kin structure among Guam rail founders: a comparison of pedigrees and DNA profiles. Molecular Ecology **3**:109–119.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series **41**:95-98.

Haltenorth T (1963) Handbuch der Zoologie Band VII: Mammalia. 1. Teil: Die Klassifikation der Saugetiere. 18. Ordnung Paarhufer Artiodactyla Owen, 1848. Walter de Gruyter & Co., München, Germany.

Hamilton-Smith C (1827) The class Mammalia. Supplement to the order Ruminantia. In: The animal kingdom, arranged in conformity with its organization, by the Baron Cuvier, with additional descriptions of all the species hitherto named, and of many not before noticed (Griffith E, Hamilton-Smith C, Pidgeon E eds.), Vol. IV. G. B. Whittaker, London, United Kingdom. Pp. 105–165

Harris LH (1971) Notes on the introduction and history of sambar deer in New Zealand. New Zealand Wildlife **35**:33–42.

Hartl GB, Zachos FE, Nadlinger K, Ratkiewicz M, Klein F, Lang G (2005) Allozyme and mitochondrial DNA analysis of French red deer (*Cervus elaphus*) populations: genetic structure and its implications for management and conservation. Mammalian Biology **70**:24–34.

Hayssen V, Van Tienhoven A, Van Tienhoven A (1993) Asdell's patterns of mammalian reproduction: a compendium of speciesspecific data. Comstock University Press, Ithaca, New York.

Heish HM Chiang HL, Tsai LC, Lai SY, Huang NE, Linacre A, Lee JC (2001) Cytochrome b gene for species identification of the conservation animals. Forensic Science International **122:**7-18**.**

Heude PM (1888) Etude sur les ruminants de l'Asia orientale. Notes sur le type Sambur, tel qu'il est en Cochinchine. Me´moires Concernant l'Histoire Naturelle de l'Empire Chinois par des Pe´res de la Compagnie de Jesus **2**:41–46.

Heusser CJ (2010) Late-glacial of southern South America. Quaternary Science Reviews **12**:345–350.

Hodgson BH (1841) Note on *Cervus elaphus* of the Sal Forest of Nepal. Hodie, *C. affinis*, nob. Journal of Asiatic Society of Bengal **10**:721–724.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs neutral genetic diversity: Implications for landscape genetics. Landscape Ecology **21**:797–807.

Holderegger R, Wagner HH (2008) Landscape Genetics. BioScience **58**:199-207.

Holsinger, Kent E, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting FST. Nature Review Genetics **10**:639–650.

Hooijer DA (1951) On the special evidence of early man in the middle Pleistocene of Southwest China. Southwestern Journal of Anthropology 7:77–81.

Hsu MJ, Agoramoorthy G (1997) Wildlife conservation in Taiwan. Conservation Biology **11**:834–836.

Hu J, Fang SG, Wan QH (2006) Genetic diversity of Chinese water deer (*Hydropotes inermis inermis*): implications for conservation. Biochemical Genetics **44**:161–172.

Huang Q-Y, Xu F-H, Shen H, Deng H-Y, Liu Y-J, Liu Y-Z, Li J-L, Recker RR, Deng H-W (2002) Mutation Patterns at Dinucleotide Microsatellite Loci in Humans. Americal Journal of Humam Genetics **70**:625634.

Huang W, Si X, Hou Y, Miller-Antonio S, Schepartz LA (1995) Excavations at Panxian Dadong, Guizhou Province, southern China. Current Anthropology **36**:844–846.

Idaghdour Y, Broderick D, Korrida A, Chbel F (2004) Mitochondrial control region diversity of the houbara bustard *Chlamydotis undulata* complex and genetic structure along the Atlantic seaboard of North Africa. Molecular Ecology **13**: 43–54.

Idris I, Moin S (2009) Somatic Chromosomes of the Bornean Sambar Deer and Rusa Deer Interspecific Hybrids. American Journal of Applied Sciences **6**: 862-868.

Isaac NJ, Turvey ST, Collen B, Waterman C, Baillie JE (2007) Mammals on the EDGE: Conservation Priorities Based on Threat and Phylogeny. PLoS ONE **2**: e296.

Jacob K (1949) Land connections between Ceylon and peninsular India. Proceeding of National Institute of Science, India **15**:341–343.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution **11**:424-429.

Jathanna D, Karanth KU, Johnsingh AJT (2003) Estimation of large herbivore densities in the tropical forests of southern India using distance sampling. Journal of Zoology (London) **261**:285 –290.

Jedrzejewski W, Niedzialkowska M, Nowak S, Jedrzejewska B (2004) Habitat variables associated with wolf (*Canis lupus*) distribution and abundance in northern Poland. Diversity and Distributions **10**:225–233

Jerdon TC (1874) The mammals of India; a natural history of all the animals known to inhabit continental India. John Wheldon, London, United Kingdom.

Johnsingh AJT (1980) Sambar, a magnificent deer. Australian Deer **5**:3–6.

Johnsingh AJT (1983) Large mammalian prey–predators in Bandipur. Journal of the Bombay Natural History Society **80**:1–57.

Johnsingh AJT, Ramesh K, Qureshi Q, David A, Goyal SP, Rawat GS *et al.* (2004) Conservation status of tiger and associated species in the Terai Arc Landscape, India. Wildlife Institute of India, Dehradun.

Johnsingh AJT, Sankar K (1991) Food plants of chital, sambar and cattle on Mundanthurai Plateau, Tamil Nadu, south India. Mammalia **55**:57–66.

Johnson WE, O'Brien SJ (1997) Phylogenetic Reconstruction of the Felidae Using 16S rRNA and NADH–5 Mitochondrial Genes. Journal of Molecular Evolution **44**:S98–S116.

Jones KC, Levine KF, Banks JD (2000) DNA-based genetic markers in black-tailed and mule deer for forensic applications. California Department of Fish and Game **86**:115–126.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology **16**: 099-1106.

Kappes SM, Keele JW, Stone RT, McGraw RA, Sonstegard TS, Smith TP, Lopez-Corrales NL, Beattie CW (1997) A second generation linkage map of the bovine genome. Genome. Res. **7**:235-249.

Karanth KU, Sunquist ME (1992) Population structure, density and biomass of large herbivores in the tropical forests of Nagarahole, India. Journal of Tropical Ecology **8**:21–35.

Karanth KU, Sunquist ME 1995. Prey selection by tiger, leopard and dhole in tropical forests. Journal of Animal Ecology **64:**439–450.

Katsares V, Tsiora A, Galinou–Mitsoudi S, Imsiridou A (2008) Genetic structure of the endangered species Pinna nobilis (Mollusca: Bivalvia) inferred from mtDNA sequences. Biologia **63**:412–417.

Kawanishi K, Rayan DM, Gumal MT, Shepherd CR (2014) Extinction process of the sambar in Peninsular Malaysia. DSG Newsletter *Nº26, ISSN 2312-4644.*

Kelton SD, Skipworth JP (1987) Food of sambar deer (*Cervus unicolor*) in a Manawatu (New Zealand) flax swamp. New Zealand Journal of Ecology **10**:149-52.

Kerr R (1792) The animal kingdom, or zoological system, of the celebrated Sir Charles Linnaeus; Class I. Mammalia In: Systema Naturae published by Geottingen G, Murray J, Faulder R, London, United Kingdom.

Khan M, Khan M (1968) Deer biological data. Malayan Nature Journal **21**:159–164.

Khun DZ, Kan LS (1991) Populations of some mammals in the sclerophylous evergreen tropic forests near Konhkanyng (South Vietnam). Zoologicheskii Zhurmal **70**:114–118.

Kim KI, Lee JH, Li K *et al.* (2002) Phylogenetic relationship of Asian and European pig breeds as determined by mtDNA D-loop sequence polymorphism. Animal Genetics **33:**19-25.

Kimura M, Crow JF (1964) The number of allele that can be maintained in a finite population. Genetics **4**:725-738.

Knowles LL, Maddison WP (2002) Statistical phylogeography. Molecular Ecology **11**:2623–2635.

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablancatt FX, Wilson AC (1989) Dynamics of Mitochondrial DNA Evolution in Animals: Amplification and Sequencing with Conserved Primers. Proceedings of the Natural Academy of Sciences, USA **86**:6196-6200.

Koizumi T, Ohtaishi N, Kaji K, Yu Y, Tokida K (1993) Conservation of white-lipped deer in China. Pp. 309–318 in *Deer in China: biology and management* (N. Ohtaishi and H-L. Sheng, eds.). Elsevier Science Publishers, Amsterdam, The Netherlands.

Kruuk LEB, Slate J, Pemberton JM, Brotherstone S, Guinness F, Clutton-Brock T (2002) Antler size in red deer: heritability and selection but no evolution. Evolution **56:**1683–1695.

Kumar A, Walker S, Mouler S (2000) Prioritization of Endangered Species. *In*: 'Setting biodiversity conservation priorities for India': summary of the findings and conclusions of the Biodiversity Conservation Prioritisation Project, Volume II edited by Singh S, World Wide Fund for Nature-India.

Kumara HN, Singh M (2004) The influence of differing hunting practices on the relative abundance of mammals in two rainforest areas of the Western Ghats, India. Oryx **38**:321–327.

Kurt F (1978) Socio-ecological organization and aspects of management of South Asian deer. Pp. 219–239 in Threatened deer: proceedings of a working meeting of the Deer Specialist Group of the Survival Service Commission. International Union for the Conservation of Nature and Natural Resources, Gland, Switzerland.

Kurten F (1968) Pleistocene mammals of Europe. Weidenfeld and Nicolson, London, United Kingdom.

Kushwaha SPS, Khan A, Habib B, Quadri A, Singh A (2004) Evaluation of sambar and muntjac habitats using geostatistical modelling. Current Science **86**:1390–1400.

Lacy RC (1997) Importance of genetic variation to the viability of mammalian populations. Journal of Mammalogy **78:**320-335.

Land DE, Lacy RC (2000) Introgression level achieved through Florida panther genetic restoration. Endangered Species Updates **17**:99–103.

Lande R (1988) Genetics and demography in biological conservation. Science **241**:1455–1460.

Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL (2007) Genetic properties influencing the evolvability of gene expression. Science **317**:118-121.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. Bioinformatics **23**:2947-2948.

Larson G, Dobney K, Albarella U *et al.* (2005) Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science **307:**1618-1621.

Lei R, Jian Z, Hu Z, Yang W (2003) Phyloge Phylogenetic relationships of Chinese antelopes (Subfamily Antilopinae) based on mitochondrial Ribosomal RNA gene sequences. Journal of Zoology **261**:227–237.

Leslie DM (2008) Boselaphus tragocamelus (Artiodactyla: Bovidae). Mammalian Species **813:**1–17.

Leslie DM (2011) *Rusa unicolor* (Artiodactyla: Cervidae). Mammalian Species **43**:1–30.

Lessa EP, D'Elia G, Pardinas UFJ (2010) Genetic footprints of late Quaternary climate change in the diversity of Patagonian-Fueguian rodents. Molecular Ecology **19**:3031–3037.

Lewis JC, Flynn LB, Marchinton RL, Shea SM, Marchinton EM (1990) Part I: Introduction, study area description and Literature review In: Ecology of Sambar deer on St Vincent National Wildlife Refuge, Florida. Bulletin of Tall Timbers Research Station **25:**1-107.

Li H, Wu J (2004) Use and misuse of landscape indices. Landscape Ecology **19**:389–399

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25:**1451-1452.

Lister AM (1987) Diversity and evolution of antler form in Quaternary deer. Pp. 81–98 in Biology and management of the Cervidae (Wemmer CM ed). Smithsonian Institution Press, Washington, D.C.

Long ES, Diefenbach DR, Rosenberry CS, Wallingford BD, Grund MD (2005) Forest cover influences dispersal distance of whitetailed deer. Journal of Mammalogy **86:**623–629.

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 (2004) Phylogeography and genetic ancestry of tigers (*Panthera tigris*). PLoS Biology **2**:2275–2293.

Lydekker R (1898) The deer of all lands: a history of the family Cervidæ living and extinct. Rowland Ward, Limited, London, United Kingdom.

Lydekker R (1915) Catalogue of the ungulate mammals in the British Museum (Natural History). Vol. IV. Artiodactyla, families Cervidæ (deer), Tragulidæ (chevrotains), Camelidæ (camels and llamas), Suidæ (pigs and peccaries), and Hippopotamidæ (hippopotamuses). Trustees of the British Museum, London, United Kingdom.

Lydekker R (1916) Wildlife of the world VolII Rowland Ward Ltd, London, UK

Lydekker R, Dollman JG (1985) The Game Animals of the Indian Subcontinent . International Books and Periodicals Supply Service, New Delhi.

Mackinnon J (2008) Order Artiodactyla. Pp. 451–480 in A guide to the mammals of China (A. T. Smith and Y. Xie, eds.). Princeton University Press, Princeton, New Jersey.

Madhusudan MD, Karanth KU (2002) Local hunting and large mammal conservation in India. Ambio **31**: 49-54.

Manel S, Schwartz K, Luikart G, Taberlet P (2003) Landscape genetics: Combining landscape ecology and population genetics. Trends in Ecology and Evolution **18**:189–197.

Mani MS (1974) Ecology and biogeography in India, Dr. W. Junk b.v. Publishers The Hague, Netherland.

Markgraf V, McGlone M, Hope G (1995) Neogene paleoenvironmental and paleoclimatic change in southern temperate ecosystems—a southern perspective. Trends in Ecology and Evolution **10**:143–147.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology **7**:639-655.

Martin C (1977) Status and ecology of the Barasingha (*Cervus duvauceli branderi*) in Kanha National Park (India). Journal of the Bombay Natural History Society **74:** 60-132.

Martínez JG, Carranza J, Fernández-García JL, Sánchez-Prieto CB (2002) Genetic variation of red deer populations under hunting exploitation in Southwestern Spain. Journal of Wildlife Management **66:**1273–1282.

Mary E, Balakrishnan M (1984) A study on olfactory signals in sambar deer, Cervus unicolor. Proceedings of the Indian Academy of Sciences **93**:71–76.

McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology **82:**290–297.

McGarigal K (2002) Landscape pattern metrics in El-Shaarawi AH, Piegorsch WW, eds Encyclopedia of Environmetrics. Pp 1135–1142.

McLoughlin PD, Paetkau D, Duda M, Boutin S (2004) Genetic diversity and relatedness of boreal caribou populations in western Canada. Biological Conservation **118**:593–598.

Merriam G (1984) Connectivity: A fundamental ecological characteristic of landscape pattern in Brandt J, Agger P, eds Proceedings of the First International Seminar of the International Association of Landscape Ecology (IALE), Organized at Roskilde University Centre, Denmark, Roskilde (Denmark) Pp 5–15.

Min MS, Okumura H, Jo DJ, An JH, Kim KS, Kim CB, Shin NS, Lee MH, Han CH, Voloshina IV, Lee H (2004) Molecular phylogenetic status of the Korean goral and Japanese serow based on partial sequences of the mitochondrial cytochrome b gene. Molecular Cells **17**:365–372.

Mitton JB (1994) Molecular approaches to population biology. Annual Reviews of Ecological Systematics **25**: 45-69.

Miyamoto MM, Kraus F, Oliver RA (1990) Phylogeny and evolution of antlered deer determined from mitochondrial DNA sequences. Proceedings of the Natural Academy of Sciences, USA **87:**6127-6131.

Mladenoff DJ, Sickley TA, Haight RG, Wydeven AP (1995) A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. Conservation Biology **9**:279–294.

Moore SS, Barendse W, Berger KT, Armitage SM, Hetzel DJS (1992) Bovine and ovine DNA microsatellites from the EMBL and GenBank databases. Animal Genetics **23**:463–467.

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetics **30**:194-200.

Morin PA, Moore JJ, Chakraborty R, Goodall J, Woodroffe DS (1994) Kin selection, social structure, gene flow and the evolution of chimpanzees. Science **265**:1193-1201.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biology **51:**238–254.

Moritz CT, Dowling E, Brown WM (1987) Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. Annual Review of Ecological Systmatics **18:**269-292.

Morris RC (1938) 'Stomping grounds' and 'sore neck' in sambar. Journal of the Bombay Natural History **40**:560–561.

Morrison JM, Sechrest W, Dinerstein E, Wilcove DS, Lamoreux JL (2007) Persistence of large mammal faunas as indicators of human impact. Journal of Mammalogy **88**:1363– 1380.

Moxon R, Willis C (1999) DNA microsatellites: Agents of evolution? Scientific American **280**:94-99.

Muir PD, Semiadi G, Asher GW, Broad TE, Tate ML, Barry TN (1997) Sambar deer (Cervus unicolor) 3 red deer (C. elaphus) interspecies hybrids. Journal of Heredity **88:**366–372.

Naidu A, Fitak RR, Munguia-Vega A, Culver M (2012) Novel primers for complete mitochondrial cytochrome b gene sequencing in mammals. Molecular Ecology Resources **12:**191–196.

Na–Nakorn U, Sukmanomon S, Nakajima M, Taniguchi N, Kamonrat W, Poompuang S, Nguyen TTT (2006) MtDNA diversity of the critically endangered Mekong giant catfish (*Pangasianodon gigas* Chevey, 1913) and closely related species: implications for conservation. Animal Conservation **9**:483–494.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**:583–590.

Ngampongsai C (19770 Habitat relations of the sambar (*Cervus unicolor*) in Khao-Yai National Park, Thailand. Ph.D. dissertation, Michigan State University, East Lansing.

Nowak RM (1999) Walker's mammals of the world. 6th ed. Johns Hopkins University Press, Baltimore, Maryland.

Nugent G, Parkes JP, Tustin KG (1987) Changes in the density and distribution of red deer and wapiti in northern Fiordland. New Zealand Journal of Ecology 10:23-33

O'Brien SJ (1994) A role for molecular genetics in biological conservation. Proceedings of the Natural Academy of Sciences, USA **91**:5748-5755.

Ohtaishi N, Gao Y (1990) A review of the distribution of all species of deer (Tragulidae, Moschidae and Cervidae) in China. Mammal Review **20**:125–144.

Onuma M, Suzuki M, Ohtaishi N (2006) Possible conservation units of the sun bear (*Helarctos malayanus*) in Sarawak based on variation of mtDNA control region. Japanese Journal of Veterinary Research **54**:135–139.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. Genetics **147**:1943-1947.

Palma RE, Rivera-Milla E, Salazar-Bravo J et al. (2005) Phylogeography of Oligoryzomys longicaudatus (Rodentia: Sigmodontinae) in temperate South America. Journal of Mammalogy **86**:191–200.

Pande HK, Arora S (2014) India's fifth national report to the Convention on Biological Diversity 2014, Ministry of Environment and Forests, Government of India, New Delhi, India.

Parker KL, Robbins CT, Hanley TA (1984) Energy expenditures for locomotion by mule deer and elk. Journal of Wildlife Management **42**:73–81.

Parson W, Pegoraro K, Niederstatter H, Foger M, Steinlechner M (2000) Species identification by means of the cytochrome b gene. International Journal of Legal Medicine **114:**23-28.

Peacock EH (1933) A game-book for Burma & adjoining territories: the types, distribution and habits of large and small game, together with notes on game preservation, photography, tracking, still-hunting and the care and measurement of trophies. H. F. & G. Witherby, Edinburgh, United Kingdom.

Pérez-Espona S, Pérez-Barbería F J, McLeod JE, Jiggins CD, Gordon IJ, Pemberton JM (2008) Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). Molecular Ecology **7**:981-996.

Pesole G, Gissi C, De Chirico A, Saccone C (1999) Nucleotide substitution rate of mammalian mitochondrial genomes. Journal of Molecular Evolution **48**:427–434.

Petronio CT, Bellucci KL, Di Stefano G (2007) Remarks on some Eurasian pliocervines: characteristics, evolution, and relationships with the tribe Cervini. Geobio **40:** 113–130

Phythian-Adams EG (1951) Jungles memories. Part IX—antelope and deer. Journal of the Bombay Natural History Society **50**:1–12.

Piggot MP, Banks SC, Taylor AC (2004) Population structure of brush-tailed rock wallaby (*Petrogale penicillata*) colonies inferred from analysis of faecal DNA. Molecular Ecology **15**:93-105.

Piggot MP, Taylor AC (2003) Remote collection of animal DNA and its application in conservation management and understanding the population biology of rare and cryptic species. Wildlife Research **30**:1-13.

Pimm SL, Pimm TL, Ayres M, Balmford A, Branch G, Brandon K (2001) Can we defy nature's end? Science **233:**2207–2208.

Pitra C, Fickela J, Meijaardb E, Groves PC (2004) Evolution and phylogeny of old world deer. Molecular Phylogenetics and Evolution **33:**880–95.

Plotka ED (1999) Deer. in Encyclopedia of reproduction (E. Knobil and J. D. Neill, eds.). Vol. 1. Academic Press, San Diego, California. Pp. 842–857.

Pocock RI (1933) The homologies between the branches of the antlers of the Cervidae based on the theory of dichotomous growth. Proceedings of Zoological Society of London **103**:377–406.

Pocock RI (1942) The skull-characters of some forms of sambar (Rusa) occurring to the east of the Bay of Bengal-Part I. Annals and Magazine of Natural History, Series 9 **11**:516–525.

Poetsch M, Seefeldt S, Maschke M, Lignitz E (2001) Analysis of microsatellite polymorphism in red deer, roe deer, and fallow deer-possible employment in forensic applications. Forensic Science International **6:**1–8.

Prasad V, Farooqui A, Tripathi S K M, Garg R, Thakur B (2009) Evidence of Late Palaeocene-Early Eocene equatorial rain forest refugia in southern Western Ghats, India. Journal of Biosciences **34**:777–797.

Price T (2008) Speciation in birds. Roberts and Co. Boulder, CO

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics **155**:945-959.

Rakotoarisoa J-E, Raheriarisena M, Goodman SM (2010) Phylogeny and species boundaries of the endemic species complex, Eliurus antsingy and E carletoni (Rodentia: Muroidea: Nesomyidae), in Madagascar using mitochondrial and nuclear DNA sequence data. Molecular Phylogenetics and Evolution **57**:11–22.

Ralls K, Ballou J (1983) Extinctions: lessons from zoos. In: Genetics and conservation: A reference for managing wild animal and plany populations: 64–

184 (Schonewald–Cox CM, Chambers SM, MacBryde B, Thomas L, Eds.). Benjamin Cummings Publishing Company, Inc., California

Randi E, Massimo P, Aleskysey D (2004) Mitochondrial DNA polymorphism in populations of Siberian and European roe deer (*Capreolus pygargus and C capreolus*). Heredity **80:**429-437.

Randi E, Mucci N, Claro-Hergueta F, Bonnet A, Douzery EJP (2001) A mitochondrial DNA control region phylogeny of the Cervinae: speciation in Cervus and implications for conservation. Animal Conservation **4**:1–11.

Ratcliffe PR (1984) Population-dynamics of red deer (*Cervus elaphus* L.) in Scottish commercial forests. *Proceedings of the Royal Society of Edinburgh.* Section B, Biological Sciences **82:**291–302.

Raufaste N, Rousset F (2001) Are partial Mantel tests adequate? Evolution **55:**1703–1705.

Raymond M, Rosset F (1995) GENEPOP (version 1.2): Population genetics software for exact test and ecumenicism. Journal of Heredity **86**:248 – 249.

Richardson WA (1972) A natural history survey of the sambar deer (Cervus unicolor) on the Powderhorn Ranch, Calhoun County, Texas. M.S. thesis, Texas A&M University, College Station.

Riley SPD, Pollinger JP, Sauvajot RM, York EC, Bromley C, Fuller TK, Wayne RK (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. Molecular Ecology **15**:1733–1741.

Roa AC, Chavarriaga-Aguirre P, Duque MC, Maya MM, Bonierbale MW, Iglesias C and Tohme J (2000) Cross species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: Allelic polymorphism and degree of relationship American Journal of Botany **87**:1647-1655.

Roach JL, Stapp P, Van Horne B, Antolin MF (2001) Genetic structure of a metapopulation of black-tailed prairie dogs. Journal of Mammalogy **82:**946–959.

Rodgers WA, Panwar HS (1988) Planning a Wildlife Protected Area Network in India A report prepared for the Ministry of Environment and Forests and Wildlife, Government of India, volumes 1 and 2.

Rodgers WA, Panwar HS, Mathur VB (2002) Wildlife Protected Area Network in India: A Review (Executive summary) Wildlife Institute of India.

Rodriguez-Serrano E, Cancino RA, Palma RE (2006) Molecular phylogeography of Abrothrix olivaceus (Rodentia: Sigmodontinae) in Chile. Journal of Mammalogy **87**:971–980.

Rohland N, Hofreiter M (2007a) Ancient DNA extraction from bones and teeth. Nature Protocol **2**:1756 – 1762.

Rohland N, Hofreiter M (2007b). Comparison and optimization of ancient DNA extraction. Biotechniques **42**:343–352.

Ronquist F (1997) Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. Systematic Biology **46**:193-201.

Rousset F (2002) Partial Mantel tests: reply to Castellano and Balleto. Evolution **56:**1874–1875.

Rull V (2009) Microrefugia. Journal of Biogeography **36**:481–484.

Russello MA, Amato G (2004) Ex situ population management in the absence of pedigree information. Molecular Ecology **13**:2829–2840.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. Nature **392**:491–494.

Saccone C, Pesole G, Sbisá E (1991) The main regulatory region of mammalian mitochondrial DNA: Structure-function model and evolutionary pattern. Journal of Molecular Evolution **33:**83-91.Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sankar K (1994) The ecology of three large sympatric herbivores (chital, sambar and nilgai) with special reference for reserve management in Sariska Tiger Reserve, Rajasthan PhD Thesis University of Rajasthan, Jaipur.

Sankar K (2008) Ungulate conservation in India. Pp. 37–50 in Wildlife biodiversity conservation (M. V. Reddy, ed.). Daya Publishing House, New Delhi, India.

Sankar K, Acharya B (2004) Sambar (*Cervus unicolor* Kerr, 1792). ENVIS Bulletin (Wildlife and Protected Areas): Ungulates of India 7, Wildlife Institute of India, Dehradun

Schaller GB (1967) The Deer and the Tiger A study of Wildlife in India The University of Chicago Press, Chicago pp 370.

Schaller GB (1998). Wildlife of the Tibetan Steppe. Chicago: University of Chicago Press. pp 383

Schipper J et al (130 authors) (2008) The status of the world's land and marine mammals: diversity, threat, and knowledge. Science **322**:225–230.

Schlötterer C, Amos B, Tautz D (1991) Conservation of polymorphic simple sequence loci in cetacean species. Nature **354:**63-65.

Schwartz M, Vissing J (2002) Paternal inheritance of mitochondrial DNA. New England Journal of Medicine **347**:576–580.

Sharma R, Stuckas H, Bhaskar R, Rajput S, Khan I, Goyal SP, Tiedemann R (2009) mtDNA indicates profound population structure in Indian tiger (*Panthera tigris tigris*). Conservation Genetics **10**: 909-914.

Shea SM, Flynn LB, Marchinton RL, Lewis JC (1990) Part II. Social behavior, movement ecology, and food habitats. Pp. 13–62 in Biology of sambar deer on St. Vincent National Wildlife Refuge, Florida. Bulletin of Tall Timbers Research Station **25:**1–107.

Sheng H, Ohtaishi N (1993) The status of deer in China. Pp. 1–11 in Deer in China: biology and management (N. Ohtaishi and H.-L. Sheng, eds.). Elsevier Science Publishers, Amsterdam, The Netherlands.

Shitara H, Hayashi JI, Takahama S, Kaneda H, Yonekawa H (1998). Maternal inheritance of mouse mtDNA in interspecific hybrids: segregation of the leaked paternal mtDNA followed by the prevention of subsequent paternal leakage. Genetics **148**:851–857.

Si X, Liu J, Zhang H, Yuan C (1993) Preliminary report on the excavation of Panxian Dadong, a Paleolithic cave-site in Guizhou Province. Acta Anthropologica Sinica **12**:113–119.

Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. Evolution **43**:1349-1368.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity **82**: 561–573.

Spear SF, Peterson CR, Matocq MD, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). Molecular Ecology **14:**2553–2564.

Steinborn R, Zakhartchenko V, Jelyazkov J, Klein D, Wolf E, Muller M, Brem G (1998) Composition of parental mitochondrial DNA in cloned bovine embryos. FEBS Letter **426**: 352–356.

Steinkellner H, Lexer C, Turetschek E and Glossl J (1997) Conservation of (GA)(n) microsatellite loci between *Quercus* species. Molecular Ecology **6**:1189-1194.

Steinmetz R, Chutipong W, Seauturien N (2006) Collaborating to conserve large mammals in Southeast Asia. Conservation Biology **20:**1391–1401.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP (2007) Putting the landscape in landscape genetics. Heredity **98**:128–142.

Sugg DW, Chesser RK, Dobson FS, Hoogland JL (1996) Population genetics meets behavioural ecology. Trends in Ecology and Evolution **11:**338–342

Summers N, Bowhunting Rusa Deer (www.bowhuntingsafari.com/SPrusadeer.html) accessed on June 2, 2014.

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in Ecology and Evolution **15:**199–203.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 40. Molecular Biology and Evolution **24**:1596-99.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution **28**:2731-2739.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7: 381–397.

Templeton AR, Routman E, Phillips CA (1995) Separating Population Structure from Population History: A Cladistic Analysis of the Geographical Distribution of Mitochondrial DNA Haplotypes in the Tiger Salamander, Ambystoma Tigrinum. Genetics **140**: 767–782.

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 (2005) Different population histories of the Mundari- and Mon-Khmer-speaking Austro-Asiatic tribes inferred from the mtDNA 9-bp deletion/insertion polymorphism in Indian populations. Human Genetics **119**:223-224

Thom WS (1937) The Malayan or Burmese sambar. Journal of the Bombay Natural History Society **39**:309–319.

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673–4680.

Timmins RJ, Evans TD (1996) Wildlife and habitat survey of the Nakai-Nam Theun National Biodiversity Conservation Area. Wildlife Conservation Society, New York, and Centre for Protected Areas and Watershed Management, Vientiane, Lao People's Democratic Republic.

Timmins RJ, Steinmetz R, Baral SH, Samba NK, Duckworth JW, Anwarul MI, Giman B, Hedges S, Lynam AJ, Fellowes J, Chan BPL, Evans T (2008) *Rusa unicolor* In: IUCN 2010 IUCN Red List of Threatened Species **Version 2010.2**. www.iucnredlist.org

Timmins RJ, Steinmetz R, Baral SH, Samba NK, Duckworth JW, Anwarul MI, Giman B, Hedges S, Lynam AJ, Fellowes J, Chan BPL, Evans T (2008) *Rusa unicolor*. The IUCN Red List of Threatened Species. **Version 2014.1**. www.iucnredlist.org accessed on 04 July 2014.

Tirosh I, Weinberger A, Carmi M, & Barkai N (2006) A genetic signature of interspecies variations in gene expression. Nature Genetics **38**:830–834.

Tobe SS, Kitchener AC, Linacre AMT (2010) Reconstructing Mammalian Phylogenies: A Detailed Comparison of the Cytochrome *b* and Cytochrome Oxidase Subunit I Mitochondrial Genes. PLoS ONE **11**:e14156.

Tóth G, Gáspári Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Research **10**:967–981.

Tungittiplakorn W, Dearden P (2002) Hunting and wildlife use in some Hmong communities in northern Thailand. Natural History Bulletin of the Siam Society **50:** 57–73.

Tun-Yin U (1967) Wild animals of Burma. Rangoon Gazette, Limited, Rangoon, Burma.

Turan C (2008) Molecular Systematic Analyses of Mediterranean Skates (Rajiformes). Turkish Journal of Zoology **32**:1–6.

Turchetto-Zolet AC, Pinheiro F, Salgueiro F, Palma-Silva C (2013) Phylogeographical patterns shed light on evolutionary process in South America Molecular Ecology **22**:1193–1213.

Turner M, Gardner RH, O'Neil RV (2001) Landscape Ecology in Theory and Practice: Patterns and Processes New York: Springer.

Vaiman D, Osta R, Mercier D, Grohs C, Leveziel H (1992) Characterization of five new bovine dinucleotide repeats. Animal Genetics **23**:537-541.

Van Bemmel ACV (19740 The concept of superspecies applied to Eurasiatic Cervidae. Zeitschrift fu¨r Sa¨ugetierkunde **38:**295–302.

Van der Kuyl AC, Ballasina DLP, Zorgdrager F (2005) Mitochondrial haplotype diversity in the tortoise species Testudo graeca from North Africa and the Middle East. BMC Evolutionary Biology **5**:29.

Varman KS, Sukumar R (1993) Ecology of sambar in Mudumalai Sanctuary, southern India. Pp. 273–284 in Deer in China: biology and management (N. Ohtaishi and H.-L. Sheng, eds.). Elsevier Science Publishers, Amsterdam, The Netherlands.

Vegas-Vilarrubia T, Rull V, Montoya E, Safont E (2011) Quaternary palaeoecology and nature conservation: a general review with examples from the neotropics. Quaternary Science Reviews **30**:2361–2388.

Verma SK, Singh L (2003) Novel universal primers establish identity of enormous number of animal species for forensic application. Molecular Ecology Notes **3**:28–31.

Vilà C, Savolainen P, Maldonado JE, Amorin IR (1997) Multiple and Ancient Origins of the Domestic Dog. Science **276**:1687–1689.

Waller JS, Servheen C (2005) Effects of transportation infrastructure on grizzly bears in Northwestern Montana. Journal of Wildlife Management **69**:985–1000.

Ward R (1896) Rowland Ward's records of big game. Rowland Ward and Company, Limited, London, United Kingdom.

Waser PM, Strobeck C (1988) Genetic signatures of inter-populations dispersal. Trends in Ecology and Evolution **13**:43-44.

Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. Evolution **38**: 1358.

White G and Powell W (1997) Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): An endangered tropical hardwood species. Molecular Ecology **6**:851-860.

White TH, Bowman JL, Leopold BD, Jacobson HA, Smith WP, Vilella FJ (2000) Influence of Mississippi alluvial valley rivers on black bear movements and dispersal: implications for Louisiana black bear recovery. Biological Conservation **95**:323–331.

Whitehead GK (1972) Deer of the world. Viking Press, New York.

Whitehead GK (1993) The Whitehead encyclopedia of deer. Swan Hill Press, Shrewsbury, United Kingdom.

Wiesner RJ, Ruegg JC, Morano I (1992) Counting target molecules by exponential polymerase chain reaction, copy number of mitochondrial DNA in rat tissues. Biochim Biophys Acta **183**:553–559.

Wu HL, Wan QH, Fang SG (2006) Population structure and gene flow among wild populations of the black muntjac (*Muntiacus crinifrons*) based on mitochondrial DNA control region sequences. Zoological Science **23**:333–340.

Zhang F, Jiang Z (2006) Mitochondrial phylogeography and genetic diversity of Tibetan gazelle (*Procapra picticaudata*): implications for conservation. Molecular Phylogenetics and Evolution **41**:313–321.

Zhao X, Li N, Guo W, X Hu, Liu Z, Gong G, Wang A, Feng J, Wu C (2004) Further evidence for paternal inheritance of mitochondrial DNA in the sheep (*Ovis aries*). Heredity **93**:399–403.

Zhdanski O (1925) Fossile Hirsche Chinas. Palaeontologia Sinica **3:**1–90.

Zink RM, Barrow-cloug GF (2008). Mitochondrial DNA under siege in avian phylogeography. Molecular Ecology **17:** 2107–2121.

APPENDICES

Preparation of Stock Solutions and Reagents

S.No.	Solution	Method of Preparation
1.	Ammonium Chloride (1 M)	Dissolved 53.49 gm of Ammonium Chloirde in 1 liter of Milli-Q water.
2.	Dye (6X)	Dissolved 0.25% Bromo Phenol Blue, 0.25% Xylene Cynol FF and Glycerol in distilled water.
3.	EDTA (0.5 M, pH 8.0)	Added 186.1 gm of Disodiumethylene diaminetetraacetate (EDTA.$2H_2O$) to 800 ml of distilled water. Stirred vigorously on a magnetic stirrer. The pH was adjusted to 8.0 with NaOH (~20 gm of NaOH pellets). Stored at 4^oC.
4.	Ethidium bromide (10 mg/ml)	1 gm of ethidium bromide was added to 100 ml distilled water. Stirred on a magnetic stirrer for several hours to ensure that the dye had dissolved. Stored at room temperature.
5.	Phenol (pH 8.0)	Liquified 600 gm of phenol at 60°C and added 0.1 % of hydroxyquinoline. To the melted phenol, added an equal volume of 1 M Tris-Cl pH 8.0. Stirred on a magnetic stirrer for 15 minutes. After the layers get separated, the upper aqueous layer was aspirated and then equal volume of 0.1 M of Tris-Cl (pH 8.0) was added and repeated as above until pH of phenolic phase was >7.8. Stored at 4^oC in Amber bottle.
6.	Phenol: Chloroform (25:24)	Mixed equal amount of phenol and chloroform. Equilibrated the mixture by extracting several time with 0.1 % Tris-Cl (pH 8.0). Stored at 4^oC in Amber bottle.
7.	Phenol: Chloroform: Isoamyl (25:24:1)	Mixed equal amount of phenol and chloroform and one volume of isoamyl alcohol. Equilibrated the

		mixture by extracting several time with 0.1 % Tris-Cl (pH 8.0). Stored at 4°C in Amber bottle.
8.	Proteinase K (20mg/ml)	Dissolved 20 mg of Proteinase K in 1ml of distilled water. Stored at –20°C
9.	Sodium Chloride (5 M)	Dissolved 292.2 gm of Sodium chloride in 1 liter of distilled water.
10.	Sodium acetate (3M, pH 5.2)	Dissolved 408.1 gm of Sodium acetate.3H2O in 800 ml of distilled water. The pH was adjusted to 5.2 with glacial acetic acid and then adjusted the final volume to 1 liter with distilled water.
11.	20% SDS or Sodium lauryl sulphate.	Dissolved 200 gm of sodium dodecylsulphate in 900 ml of distilled water. Heated to 60°C to assisted dissolution. The pH was adjusted to 7.2 by adding few drops of concentrated HCl. The volume was adjusted to 1 liter with distilled water.
12.	Tris-HCl (1 M)	Dissolved 121.14 gm of Tris base in 800 ml of distilled water. The pH was adjusted to 8.0 by adding 42 ml of concentrated HCl. The volume of solution was adjusted to 1 liter with distilled water.
13.	50X TAE	Dissolved 242 gm of Tris base and 37.2 gm of EDTA.2H2O in 900 ml deionized water. Added 57.1 ml glacial acetic acid and the final volume were adjusted to 1 liter. Stored at 4°C.
14.	TE (pH 8.0)	Mixed 10 mM of Tris-Cl (pH 8.0) and 1 mM EDTA and makeup the final volume upto 50 ml.

Preparation of Lysis Buffer for Phenol: Chloroform based DNA extraction

S.No.	Chemical Name	Working concentration	Storage Temp.
1.	Tris-HCl (1 M)	50 mM	Room temp.
2.	EDTA (0.5M) (pH 8.0)	10 mM	4°C
3.	NaCl	150 mM	Room temp.

Agarose Gel Electrophoresis

S.No.	Chemical Name	Stock	Working	Storage Temp.
1.	Agarose powder (Himedia, Sigma)	100 gm	0.8% (Genomic DNA), 2% (PCR product)	Room temp.
2.	TAE Buffer (Sigma)	50X	1X	4°C
3.	Ethidium Bromide (Sigma)	10 mg/ml	0.5 g/ml	4°C
4.	Loading Buffer (Qiagen)	6X	1X	4°C
5.	DNA Ladder (Qiagen)	100bp	100bp	4°C

Polymerase Chain Reaction

S. No.	Chemical Name	Stock	Working	Storage Temp.
1.	Dream Taq DNA Polymerase (Fermentas)	5 Unit/l	0.5 Unit	20°C
2.	Di-deoxyribose nucleotides (Fermentas)	10 mM	200 M each (dNTP)	20°C
3.	Taq Buffer (Fermentas)	10X	1X	20°C
4.	Primers (Life Technologies)	10 nmol	4 pmol	20°C
5.	Magnesiumchloride (Fermentas)	25 mM	2 mM	20°C

LIST OF PUBLICATIONS

1. **Gupta SK**, Kumar A, Hussain SA **(2014)** Novel primers for sequencing of the complete mitochondrial cytochrome b gene of ungulates using non-invasive and degraded biological samples. *Conservation Genetics Resources* ***In press*** (DOI: 10.1007/s12686-014-0143-6).

2. **Gupta SK**, Kumar A, Hussain SA, Vipin, Singh L **(2013)** Cytochrome b based genetic differentiation of Indian wild pig (*Sus scrofa cristatus*) and domestic pig (*Sus scrofa domestica*) and its use in wildlife forensics. *Science and Justice* **53**: 220-222.

3. **Gupta SK**, Kumar A, Hussain SA **(2013)** Extraction of PCR-amplifiable DNA from a variety of biological samples with uniform success rate. *Conservation Genetics Resources* **5**: 215-217.

Conservation Genet Resour
DOI 10.1007/s12686-014-0143-6

TECHNICAL NOTE

Novel primers for sequencing of the complete mitochondrial cytochrome *b* gene of ungulates using non-invasive and degraded biological samples

Sandeep Kumar Gupta · Ajit Kumar · Syed Aniul Hussain

Received: 12 September 2012 / Accepted: 9 January 2014

Abstract We describe a set of novel primers for successful amplification of the complete mitochondrial cytochrome *b* (cyt *b*) gene of ungulate species. DNA extracted from non-invasively obtained and decomposed samples is found to be degraded and inappropriate for amplification of the complete gene (more than 1 kb) in single PCR amplification. We developed a series of six ungulate-specific conserved primers for the cyt *b* gene. These primers, in various combinations, amplify 366–1,266 bp fragments. We also used them in multiplex PCR reaction to assess the maximum possible size of the PCR amplification product.

Keywords Mitochondrial cytochrome *b* gene · Non-invasive and decomposed samples · PCR amplification · Ungulate species · Multiplex PCR

Introduction

DNA sequencing-based species identification relies on the amount and veracity of the sequence data available at the appropriate database. Identification of species from a sample of unknown origin is based on comparison of DNA sequences obtained from the sample against existing sequences in databases (Verma and Singh 2003; Gupta et al. 2005). A lack of reference sequences from the matching species in the database may lead to high-scoring matches with closely related sequences and produce unreliable results (Naidu et al. 2012). The International Society of Forensic Genetics (ISFG) recommends the use of an in-house and authentic DNA database for species identification for forensic validation (Linacre et al. 2011). Mitochondrial DNA (mtDNA) cytochrome *b* (cyt *b*) is an appropriate molecular marker for species identification (Tobe et al. 2010). Hence, large amounts of reference sequence data for the complete cyt *b* gene are strongly needed for phylogenetic and forensic studies. The use of a conserved primer will be extremely useful for generating a database.

The mitochondrial genome is a widely preferred locus for the development of conserved primers (Kocher et al. 1989; Verma and Singh 2003). Such primers have proved to be valuable for species identification in the wildlife forensics (Gupta et al. 2005). Recently, a pair of conserved primer for the complete cyt *b* gene was developed for mammals (Naidu et al. 2012). Amplifying the complete cyt *b* gene requires high-quality DNA; however, conservation and forensic geneticists often depend upon non-invasively obtained and forensic (mostly degraded) biological samples. Amplification and analysis of longer fragments (more than 1 kb) from DNA extracted from such samples is challenging. A primer amplifying shorter fragments will be more useful for such studies. We are working on the genetics of ungulates. Hence, we describe novel primers for amplification of the complete mtDNA cyt *b* gene of these animals from degraded samples.

Complete cyt *b* sequences of various ungulate species were obtained from GenBank and aligned using ClustalW (Thompson et al. 1994). Three forward and three reverse primers were designed on the basis of sequence homology (Table 1) and procured (Sigma Life Sciences, India).

Biological samples received for species identification in cases of ungulate poaching and fecal pellets of deer species were used. DNA was extracted (Gupta et al. 2013) and

S. K. Gupta (✉) · A. Kumar · S. A. Hussain
Wildlife Institute of India, Chandrabani,
Dehra Dun 248 001, UK, India
e-mail: skg@wii.gov.in; skg.bio@gmail.com

Published online: 19 January 2014

Table 1 Six primers designed for amplification of either complete or partial fragment of mtDNA cyt *b* gene

Primer position is derived from the complete mtDNA sequence of *Rusa unicolor swinhoei* (EF035448)

S. no.	Primer name	Sequence (5′ → 3′)
1	MC *b* (14081-105) F1	CATTATTCTCACATGGAATCTAACC
2	MC *b* (14560-584) F2	GAGGACAAATATCATTCTGAGGAGC
3	MC *b* (14584-560) R1	GCTCCTCAGAATGATATTTGTCCTC
4	MC *b* (14981-934) F3	ACCCAGACAACTACACCCCAGCAA
5	MC *b* (14934-981) R2	TTACTGGGGTGTAGTTGTCTGGGT
6	MC *b* (15347-323) R3	CTCCTTTTCTGGTTTACAAGACCAG

Table 2 All possible combinations of the primers during PCR amplification and expected amplicon lengths

S. no.	Primer combination	Expected amplicon length(s) (bp)
1	MC *b* (14081-105) F1 and MC *b* (14584-560) R1	503
2	MC *b* (14081-105) F1 and MC *b* (14934-981) R2	853
3	MC *b* (14081-105) F1 and MC *b* (15347-323) R3	1,266
4	MC *b* (14560-584) F2 and MC *b* (14934-981) R2	374
5	MC *b* (14560-584) F2 and MC *b* (15347-323) R3	787
6	MC *b* (14981-934) F3 and MC *b* (15347-323) R3	366
In multiplex reactions		
7	F1, R1, R2 and R3	503, 853 and 1,266
8	F2, R2 and R3	374 and 887
9	F1, F2, R2 and R3	374, 787, 853 and 1,266
10	F1, F2, F3 and R3	366, 787 and 1,266

Table 3 List of tested species for validation of primer

S. no.	Amplification and sequencing with ungulate species	Result	No. of individuals
1	Sambar (*Rusa unicolor*)	Successful	9
2	Swamp deer (Rucervus duvaucelii)	Successful	2
3	Chital (*Axis axis*)	Successful	3
4	Hog deer (*Axis porcinus*)	Successful	3
5	Barking deer (*Muntiacus muntjak*)	Successful	2
6	Black buck (*Antelope cervicapra*)	Successful	2
7	Indian gazelle (*Gazella bennettii*)	Successful	2
8	Nilgai (*Boselaphus tragocamelus*)	Successful	4
9	Indian wild pig (*Sus scrofa*)	Successful	8
10	Domestic pig (*Sus scrofa*)	Successful	6
Amplification with non ungulate species			
11	Tiger (*Panthera tigris*)	No amplification	
12	Leopard (*Panthera pardus*)	No amplification	
13	Human (*Homo sapiens*)	No amplification	
14	Common langur (*Semnopithecus entellus*)	No amplification	
15	Palm civet (*Viverricula indica*)	No amplification	
16	Different fish species	No amplification	

subjected to PCR amplification using the primer combinations listed in Table 2. The amplification was carried out in a 20 µl reaction volume containing 1 µl of the extracted DNA, 100 µM of dNTPs, 4 pmol of each primer, 1.5 mM $MgCl_2$, 0.5 units of AmpliTaq Gold (Invitrogen Life Technologies), and 1 × PCR buffer (10 mM Tris–HCl, pH 8.3, and 50 mM KCl). The PCR conditions were the following: initial denaturation at 95 °C for 10 min, followed by 35 cycles of denaturation at 95 °C for 45 s, annealing at 55 °C for 1 min, and extension at 72 °C for 1.5 min. The final extension was at 72 °C for 10 min. The PCR products were electrophoresed on 2 % agarose gel, stained with

ethidium bromide (0.5 mg/ml), and visualized under a UV transilluminator.

The PCR products obtained were sequenced (ABI 3130, Genetic Analyser) on both strands. The sequences were aligned using SeqScape v2.5 (Applied Biosystems) and blasted using the program BLAST to confirm the origin of the sample (Altschul et al. 1997).

Results and conclusion

PCR amplicons of expected length were obtained with different primer combinations (Table 2). Shorter fragments of length 366, 374, and 503 bp could be amplified with the DNA from all the biological samples, including fecal pellets (Table 2). A PCR product of length 1,266 bp was amplified with the DNA extracted from fresh and less decomposed samples of flesh. However, in the majority of moderately putrefied flesh samples, amplicons of length 853 and 787 bp were obtained. After alignment of the DNA sequences obtained from all the above amplicons, complete gene sequences were obtained for various ungulate species. A complete list of the species tested and the corresponding results are provided in Table 3.

Enrichment of DNA sequences in the database of forensic laboratories and GenBank is one of the essential steps for the success of forensic work and identification of species with precision. For successful comparison of sequences of unknown biological samples, either suitable fragments or complete cyt *b* gene sequences from various animal species are essential. Obtaining permissions for collecting biological samples invasively from free-ranging animals of endangered and protected species is a challenging task. However, collection of non-invasive samples is an efficient way of increasing the sample size. Hence, amplification of smaller fragments and alignment of sequences is a convenient way of enhancing databases for species for which samples are difficult to obtain. The availability of novel and complete gene sequences can improve the reliability of sequence-based species identification. The primers described here will be beneficial for the amplification and sequencing of the complete cyt *b* gene of various ungulate species using non-invasively obtained biological samples.

Acknowledgments This study was funded by the grant-in-aid of the WII. We acknowledge support from the Director, Dean (FWS), FTO and Nodal Officer (Forensic Cell) of the WII.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman J (1997) Gapped blast and psi-blast: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Gupta SK, Verma SK, Singh L (2005) Molecular insight into a wildlife crime: the case of a peafowl slaughter. Forensic Sci Int 154:214–217

Gupta SK, Kumar A, Hussain SA (2013) Extraction of PCR-amplifiable DNA from a variety of biological samples with uniform success rate. Conserv Genet Resour 5:215–217

Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. PNAS USA 86:6196–6200

Linacre A, Gusmao L, Hecht W, Hellmann AP, Mayr WR, Parson W, Prinz M, Schneider PM, Morling N (2011) ISFG: recommendations regarding the use of non-human (animal) DNA in forensic genetic investigations. Forensic Sci Int Genet 5:501–505

Naidu A, Fitak RR, Munguia-Vega A, Culver M (2012) Novel primers for complete mitochondrial cytochrome *b* gene sequencing in mammals. Mol Ecol Resour 12:191–196

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tobe SS, Kitchener AC, Linacre AMT (2010) Reconstructing mammalian phylogenies: a detailed comparison of the cytochrome *b* and cytochrome oxidase subunit I mitochondrial genes. PLoS One 11:e14156

Verma SK, Singh L (2003) Novel universal primers establish identity of enormous number of animal species for forensic application. Mol Ecol Notes 3:28–31

Science and Justice 53 (2013) 220–222

Contents lists available at SciVerse ScienceDirect

Science and Justice

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

Technical note

Cytochrome *b* based genetic differentiation of Indian wild pig (*Sus scrofa cristatus*) and domestic pig (*Sus scrofa domestica*) and its use in wildlife forensics

Sandeep Kumar Gupta [a,*], Ajit Kumar [a], Syed Ainul Hussain [a], Vipin [a], Lalji Singh [b]

[a] *Wildlife Institute of India, Dehradun, India*
[b] *Centre for Cellular and Molecular Biology, Hyderabad, India*

ARTICLE INFO

Article history:
Received 27 July 2012
Received in revised form 21 September 2012
Accepted 27 September 2012

Keywords:
Indian wild pig
Domestic pig
Wildlife crime
Genetic variation
Cytochrome *b* gene
DNA sequence variation

ABSTRACT

The Indian wild pig (*Sus scrofa cristatus*) is a protected species and listed in the Indian Wildlife (Protection) Act, 1972. The wild pig is often hunted illegally and sold in market as meat warranting punishment under law. To avoid confusion in identification of these two subspecies during wildlife forensic examinations, we describe genetic differentiation of Indian wild and domestic pigs using a molecular technique. Analysis of sequence generated from the partial fragment (421 bp) of mitochondrial DNA (mtDNA) cytochrome *b* (Cyt *b*) gene exhibited unambiguous (>3%) genetic variation between Indian wild and domestic pigs. We observed nine forensically informative nucleotide sequence (FINS) variations between Indian wild and domestic pigs. The overall genetic variation described in this study is helpful in forensic identification of the biological samples of wild and domestic pigs. It also helped in differentiating the Indian wild pig from other wild pig races. This study indicates that domestic pigs in India are not descendent of the Indian wild pig, however; they are closer to the other wild pig races found in Asia and Europe.

1. Introduction

The wild pig (*Sus scrofa*) is a widely distributed species of Europe, Asia and North Africa belonging to the Class Mammalia, Family Suidae. It is a common game animal and hence extensively exploited for sport and meat. They are highly vulnerable to human persecution due to their tendency to raid crops. Indian wild pig is a separate subspecies (*S. s. cristatus*). Though the World Conservation Union (IUCN) considers wild pig as 'Least Concern' species [1], it is a protected species in many countries. The wild pig is a protected species under Schedule-III of Indian Wildlife (Protection) Act, 1972. The domestic pig (*S. s. domestica*) is the other subspecies that is a source of pork and is an important farm animal. In the recent past a few genetic studies have been carried out to distinguish the wild and domestic pigs in Japan and China [2–5]. Despite a need to develop protocol for unambiguous identification of these two subspecies, no study has investigated the differentiation of biological samples of wild and domestic pigs in India. Since the Indian wild pig is extensively poached for meat, an unambiguous differentiation protocol for Indian wild and domestic pig can be used for effective implementation of the Wildlife (Protection) Act, 1972 so as to control wildlife crimes.

The DNA sequence variation in mtDNA genes [6–8] has proved an asset for identifying the species of confiscated biological samples [9]. In this study the partial sequence of the Cyt *b* gene [8] of the Indian wild pig and that of the domestic pig were compared. The Cyt *b* sequences for other wild pig races were obtained from GenBank through the NCBI website (http://www.ncbi.nlm.nih.gov) and compared with the Cyt *b* sequences of Indian wild and domestic pigs.

2. Material and methods

2.1. DNA isolation

We used 42 and 31 tissue samples of Indian wild and domestic pigs, respectively collected from various parts of India for characterization of the DNA sequence. DNA was extracted from these samples using the phenol/chloroform method [10] and DNeasy Blood Tissue Kit (QIAGEN, Germany) in a final elution volume of 40–100 μL.

2.2. PCR amplification and DNA sequencing

The DNA was subjected to PCR amplification by using universal primers [8]. PCR amplification was performed in a GeneAmp PCR System 2700 (Applied Biosystems, Singapore) in a final volume of 20 μL containing 10–40 ng of extracted DNA, 1×PCR buffer (Applied Biosystem), 2.0 mM $MgCl_2$, 0.2 mM of each dNTP, 3 pmol of each primer, and 0.5 units of AmpliTaq Gold DNA polymerase (Applied Biosystems). The amplification conditions were the following: 95 °C for 10 min, followed by 35 cycles at 95 °C for 45 s, 55 °C for 45 s and 72 °C for 1 min, with a final extension of 72 °C for 10 min. The

* Corresponding author at: Wildlife Institute of India, Chandrabani, Dehra Dun, 248 001 (U.K.), India. Tel.: +91 135 2640111 15x353; fax: +91 135 2640117.
E-mail address: skg@wii.gov.in (S.K. Gupta).

1355-0306/$ – see front matter
http://dx.doi.org/10.1016/j.scijus.2012.09.005

PCR products were electrophoresed on 2% agarose gel and visualized under UV light in the presence of ethedium bromide dye.

The PCR products obtained were treated with Exonuclease *I* (Exo *I*) and Srimp Alkaline Phosphatase (SAP) to clean the unused primers and dNTPs. 1.5 μL of each PCR product was directly incubated with 0.5 μl ExoSAP-IT (USB, Cleveland, Ohio) at 37 °C for 20 min. followed by inactivation of enzymes at 85 °C for 15 min. The cleaned PCR products were sequenced (with Applied Biosystems Genetic Analyzer) using BigDye 3.1 sequencing kit (Applied Biosystems) from both strands. The sequences were aligned and cleaned by using Sequencher 4.7 (Gene Code Corporation).

2.3. Sequence analysis

The Cyt *b* sequences obtained from the Indian wild and domestic pig samples were aligned using ClustalW [11] along with all the suidae sequences obtained from GenBank. The Hasegawa–Kishino–Yano (HKY + G) using a discrete Gamma distribution model [12] has the lowest Bayesian Information Criterion (BIC) score among all the models tested, using MEGA 5 [13]. Hence, it is considered the best model for nucleotide substitution. The aligned sequences were used for construction of Neighbour-Joining (NJ) trees (Fig. 1) at bootstrapping for 1000 replications using Kimura-2 parameter by MEGA 5 [13].

3. Results

All the 42 Cyt *b* sequences generated from Indian wild pigs matched with 99% homology. Based on the sequence similarity, the DNA sequences of all the Indian wild pigs were divided into three groups and coded as IWB-1, 2 and 3 (Table 1). Of the 42 sequences, 20 were similar and grouped as IWB-1; 12 sequences were grouped as IWB-2; and 10 sequences were grouped as IWB-3. In comparison with IWB-1, IWB-2 has a substitution of G in place of A at position 580 and IWB-3 has substitutions of T in place of C at positions 695 and 770 (Table 1). The number of nucleotide positions was generated from the complete Cyt *b* gene of Lanyu wild boar (acc. no. EF375877). The sequences of IWB-1, 2 and 3 were submitted to GenBank with acc. no. JN039027, JN039028, and JN039029, respectively.

The Cyt *b* gene sequences generated for 31 domestic pigs collected from different parts of India matched with 99% homology. Based on sequence homology, three distinct groups of domestic pigs were identified as DP-1, DP-2 and DP-3, which consisted of 11, 18 and 2 individuals, respectively. The sequences of DP-1, 2 and 3 were submitted to GenBank with acc. no. JN039030, JN039031 and JX411960, respectively.

The Cyt *b* sequences of the 42 Indian wild pig samples exhibited only 97% homology with those of the domestic pigs. The variable nucleotide positions of the Cyt *b* sequences from Indian wild and domestic pigs are shown along with variable nucleotide positions of all other wild pig races of the world in Table 1. The NJ tree demonstrated that all three groups of Indian wild pig are restricted to one cluster that significantly differed from the clusters of all other wild pigs of the world and domestic pigs of India (Fig. 1). The 3% genetic variation within one of the evolutionary conserved genes between the Indian wild and domestic pigs and the other wild pig races separated the Indian wild pigs from the other two forms. Unambiguous nine forensically informative nucleotide variations were observed between Indian wild and domestic pigs (Table 1), which are the basis for differentiation of the forensic samples with precision.

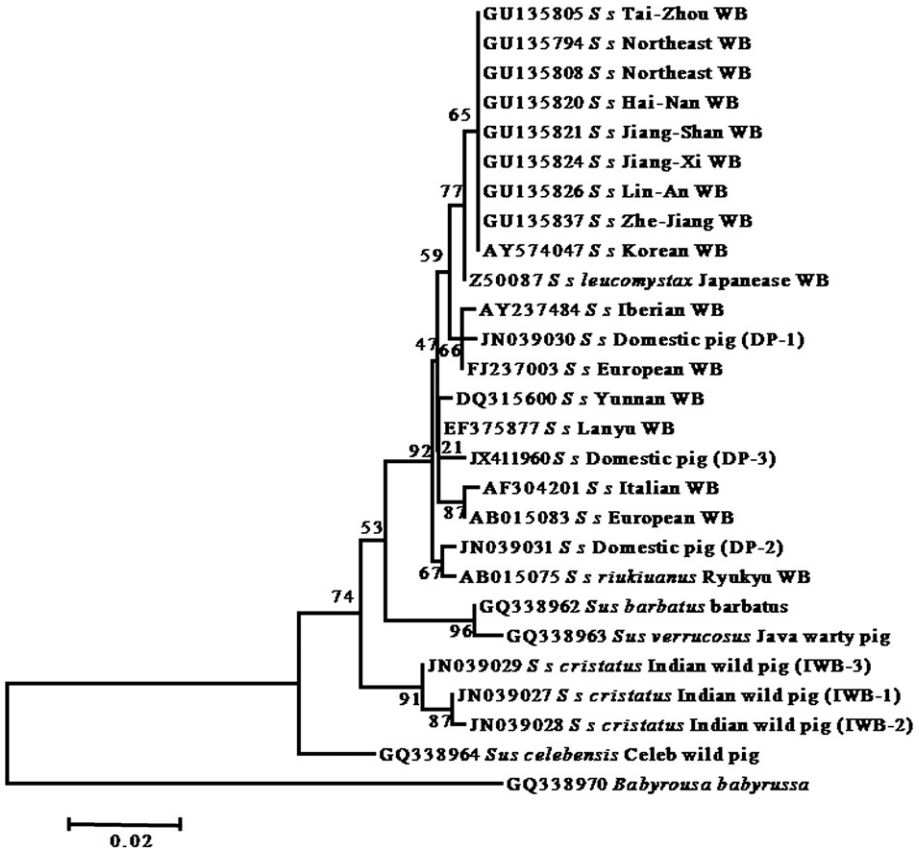


Fig. 1. The Neighbour-Joining (NJ) tree constructed on the basis of partial fragment of Cyt *b* gene demonstrating the phylogenetic relationship of wild pig(s) and domestic pig. The IWB-1, IWB-2 and IWB-3 are the three distinct haplotypes of Indian wild pig. The DP-1, DP-2 and DP-3 are the three distinct haplotypes of Indian domestic pig. The name of the wild and domestic pigs is given after the GenBank (NCBI) accession number.

Table 1
Sequence variation in partial fragment (421 bp) of Cyt *b* gene generated by universal primer [8]. The examined region expands between 423 and 843 nucleotide positions of a complete Cyt *b* gene of a wild boar (acc. no. EF375877). Three digit numeric values at top is the position of variable nucleotide. "." for indicating "similarity" for all identical nucleotides. Alpha-numeric in the brackets () are the NCBI GenBank accession number of DNA sequence.

	4 4 4	4 6 2	4 6 6	4 9 8	5 4 3	5 5 1	5 5 5	5 5 8	5 8 0	5 8 4	5 9 8	6 1 8	6 9 3	6 9 5	7 0 0	7 2 2	7 2 6	7 3 0	7 3 6	7 5 3	7 6 8	7 7 0	7 7 6	8 3 7	8 4 0	8 4 1
S. scrofa cristatus -IWB-1 (JN039027)	T	C	G	A	C	T	G	C	G	T	T	T	A	T	C	C	A	C	T	A	C	T	C	C	T	C
S. scrofa cristatus -IWB-2 (JN039028)	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S. scrofa cristatus -IWB-3 (JN039029)	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	C	.	.	.	.
S.scrofa domestica-DP-1 (JN039030)	.	T	A	G	.	.	.	A	.	.	C	C	.	C	T	T	T	.	.	.	.	C	.	T	.	G
S.scrofa domestica-DP-2 (JN039031)	.	T	A	G	T	.	.	A	.	.	C	.	.	C	T	T	C	A	.	.	.	C	.	T	.	.
S.scrofa domestica-DP-3 (JX411960)	.	T	A	G	T	.	.	A	.	.	C	.	.	C	T	T	T	.	A	.	.	C	T	T	.	.
Ynnan wild boar (DQ315600)	C	T	A	G	T	.	.	A	.	.	C	.	.	C	T	T	T	.	.	.	.	C	.	T	.	.
Lanyu wild boar (EF375877)	.	T	A	G	T	.	.	A	.	.	C	.	.	C	T	T	T	.	.	.	.	C	.	T	.	.
Ryukyu wild boar (AB015075)	.	T	A	G	T	C	.	A	.	.	C	.	.	C	T	T	C	.	.	.	.	C	.	T	.	.
European wild boar (FJ237003)	.	T	A	G	.	.	.	A	.	.	C	C	.	C	T	T	T	.	.	.	.	C	.	T	.	.
Italian wild boar (AF304201)	.	T	A	G	T	.	.	A	.	G	C	.	.	C	T	T	T	.	.	G	T	C	.	T	.	.
Korean wild boar (AY574047)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	C	.
European wild boar (AB015083)	.	T	A	G	T	.	.	A	.	.	C	.	.	C	T	T	T	.	.	G	T	C	.	T	.	.
Zhe-Jiang wild boar (GU135837)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	C	.
Lin-An wild boar (GU135826)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Jiang-Xi wild boar (GU135824)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Jiang-Shan wild boar (GU135821)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Hai-Nan wild boar (GU135820)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Northeast wild boar (GU135808)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Tai-Zhou wild boar (GU135805)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Northeast wild boar (GU135794)	.	T	A	G	T	.	.	A	.	.	C	C	G	C	T	T	T	.	.	.	.	C	.	T	.	.
Japanese wild boar (Z50087)	.	T	A	G	T	.	.	A	.	.	C	C	.	C	T	T	T	.	.	.	.	C	.	T	.	.
Iberian wild boar (AY237484)	.	T	A	G	.	.	A	A	.	.	C	C	.	C	T	T	T	.	.	.	.	C	.	T	.	.

4. Discussion and conclusions

In comparison with other phylogenetic markers, the Cyt *b* gene demonstrates greater congruence with conventional mammalian phylogeny and shows a greater level of nucleotide variation in shorter sequences [14]. A partial sequence of the Cyt *b* gene demonstrates significant variations. It allowed this gene to be used as a useful marker in unambiguous identification of the Indian peafowl in the investigation of a wildlife crime [9]. In this study, the same region of the Cyt *b* gene was used to detect unambiguous alliance of domestic pig and Indian wild pigs along with the wild pigs of the world. Our study indicates that the DNA sequence generated for 421 bp fragment of Cyt *b* gene from the suspected biological samples of Indian wild pig or domestic pig can be applied for differentiation of these two subspecies. This approach for distinguishing the biological samples of Indian wild and domestic pig has been an asset in examination of more than 50 wildlife offences cases. Hence, this can be used by other forensic laboratories for differentiating the case samples of Indian wild and domestic pigs on routine basis. Moreover, in a close observation of the NJ tree (Fig. 1), it is evidenced that domestic pigs in India are not descendent of the Indian wild pigs, which are 99% similar with the other wild pig races of Europe and Asia (Table 1).

Acknowledgements

This work is supported from the grants from Department of Biotechnology, Central Zoo Authority and Ministry of Environment and Forest, Government of India through Grant-in-aid to WII. We are thankful to Mr PR Sinha, Director; Dr. VB Mathur, Dean and Dr. SP Goyal, Nodal Officer, Forensics Cell from WII for their support during a part of this study in WII. We also thank to Dr. S. Shivaji from CCMB for his generous help during a part of this work at CCMB. We sincerely acknowledge the help of Forest Department from several Indian states for forwarding the wildlife offence cases and known samples of Indian wild pig for analysis.

References

[1] W. Oliver, K. Leus, Sus scrofa, in: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.1 (www.iucnredlist.org) (2008) access on 07 June 2012.
[2] N. Ishiguro, Y. Naya, M. Horiuchi, M. Shinagawa, A genetic method to distinguish crossbred inobuta from Japanese wild boars, Zoological Science 19 (2002) 1313–1319.
[3] H. Hongo, N. Ishiguro, T. Watanobe, N. Shegehara, T. Anezaki, V.T. Long, D.V. Binh, N.T. Tien, N.H. Nam, Variation in mitochondrial DNA of Veietnamese pigs: relationship with Asian domestic pigs and Ryukyu wild boars, Zoological Science 19 (2002) 1329–1335.
[4] T. Watanobe, N. Ishiguro, M. Nakano, Phylogeography and population structure of the Japanese wild boar *Sus scrofa leucomystax*: mitochondrial DNA variation, Zoological Science 22 (2003) 1477–1489.
[5] T. Watanobe, N. Ishiguro, M. Nakano, A. Matsui, H. Hongo, K. Yamazaki, O. Takahashi, Prehistoric Sado island populations of *Sus scrofa* distinguished from contemporary Japanese wild boar by ancient mitochondrial DNA, Zoological Science 21 (2004) 219–228.
[6] W. Parson, K. Pegoraro, H. Neiderstatter, M. Foger, M. Steinlechner, Species identification by means of cytochrome *b* gene, International Journal of Legal Medicine 114 (2000) 23–28.
[7] H.M. Hsieh, L.H. Huang, L.C. Tsai, Y.C. Kuo, H.H. Meng, A. Linacre, J.C.I. Lee, Species identification of rhinoceros horns using the Cyt *b* gene, Forensic Science International 136 (2003) 1–11.
[8] S.K. Verma, L. Singh, Novel universal primers establish identity of enormous number of animal species for forensic application, Molecular Ecology Notes 3 (2003) 28–31.
[9] S.K. Gupta, S.K. Verma, L. Singh, Molecular insight into a wildlife crime: the case of a peafowl slaughter, Forensic Science International 154 (2005) 214–217.
[10] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular cloning: a laboratory manual, second ed. Cold Spring Harbor Press, New York, NY, 1989.
[11] J.D. Thompson, D.G. Higgins, T.J. Gibson, Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Research 22 (1994) 4673–4680.
[12] M. Hasegawa, H. Kishino, T. Yano, Dating of the human-ape splitting by a molecular clock of mitochondrial DNA, Journal of Molecular Evolution 22 (1985) 160–174.
[13] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Molecular Biology and Evolution 28 (2011) 2731–2739.
[14] S.S. Tobe, A.C. Kitchener, A.M.T. Linacre, Reconstructing mammalian phylogenies: a detailed comparison of the cytochrome *b* and cytochrome oxidase subunit I mitochondrial genes, PloS One 11 (2010) e14156.

Conservation Genet Resour (2013) 5:215–217
DOI 10.1007/s12686-012-9772-9

TECHNICAL NOTE

Extraction of PCR-amplifiable DNA from a variety of biological samples with uniform success rate

Sandeep Kumar Gupta · Ajit Kumar · Syed Ainul Hussain

Received: 1 March 2012 / Accepted: 29 August 2012 / Published online: 11 September 2012

Abstract This study describes the comparison of three DNA extraction protocols for successful extraction of PCR amplifiable quality DNA from bones, antlers and feces samples of Sambar deer (*Rusa unicolor*). Three different DNA extraction protocols were compared in this study including Phenol–Chloroform (PC), column based Qiagen kit, and Guanidine hydrochloride (Gu-HCl) based in-house method. The effectiveness of the protocols was compared for higher success rate of PCR amplification from the extracted DNA. This study highlights that silica based indigenous DNA extraction protocol using Gu-HCl chaotropic salts yields better quality DNA with higher PCR amplification success rate.

Keywords DNA extraction protocol · PCR · Non-invasive wildlife samples · PCR inhibitors

Introduction

The extraction of DNA from variety of biological samples is the first and important step in the field of molecular genetics. In the field of molecular genetics, the quantity and quality of extracted DNA and the success rate of PCR amplification with extracted DNA has great importance (Gupta et al. 2011). DNA extraction from intricate biological samples including bone, antlers and fecal matters with higher PCR success rate is a challenging task. The DNA extracted from these samples contains inhibitors which affect the PCR success (Rohland and Hofreiter 2007a, b). The conundrum in selecting the better DNA extraction protocol with higher PCR success rate is one of the frequently arising situations in conservation genetics. The chaotropic salts [Guanidine thiocyanate (Gu-SCN)] and silica binding based DNA extraction technique has been used for the extraction of DNA from hard tissue including bone and teeth (Höss and Pääbo 1993; Rohland and Hofreiter 2007a), and feces (Wehausen et al. 2004). The eluted DNA contains minimum inhibitory effect but these protocols are based on Gu-SCN, which is an expensive and hazardous reagent. This study highlights the use of more efficient, less toxic and low-cost Guanidine hydrochloride (Gu-HCl).

We are conducting research to address the phylogeny of Sambar deer (*Rusa unicolor*) populations across India. As a part of this work we have to standardize the protocol for the extraction of good quality DNA from various biological samples of Sambar deer. The antler, tissue and bone samples of Sambar deer collected by forest department from samples confiscated for forensics investigation were used for optimizing the DNA extraction. The antlers and bone were stored at room temperature and tissues were stored at −20 °C. The feces of Sambar deer were collected in 70 % ethanol and stored at room temperature. For uniformity and comparability of the results from the tested protocols, same samples in equal quantity were used for the extraction of DNA in equal final volume (80 µl). For the experiment, 1.5 g of bone and antler were pulverized and incubated with 0.5 M EDTA for 48 h for decalcification. 0.5 g of the surface layer of feces was taken for each extraction. The three DNA extraction methods used in this study are Phenol–Chloroform (PC) (Sambrook et al. 1989), QIAamp DNA Stool Mini/blood and tissue kit (Qiagen, Germany), and Gu-HCl based silica binding protocol. For Gu-HCl based DNA extraction, 0.5 g of above sample was mixed

S. K. Gupta (✉) · A. Kumar · S. A. Hussain
Wildlife Institute of India, Post Box # 18, Chandrabani, Dehra Dun 248 001, Uttarakhand, India
e-mail: skg@wii.gov.in; skg.bio@gmail.com

with 500 µl of lysis buffer (6 M Gu-HCl) and 20 µl of Proteinase-K (from 20 mg/ml stock) in a 2 ml centrifuge tube and incubated at 56 °C in an hybridization oven with continuous rotation for 24–48 h (till complete lysis). The tubes were centrifuged at 13,000 rpm for 2 min and clear solutions were transferred to a fresh tube. 30 µl of silica suspension (SiO_2 powder in equal volume of distilled water) was added and incubated with constant rotation at room temp for 15 min. The tubes were centrifuged at 13,000 rpm for 2 min and the supernatant were discarded. For washing the DNA–silica pellet, 500 µl of wash buffer (20 mM Tris–HCl, pH 7.8, 1 mM EDTA, 50 mM NaCl, 50 % ethanol) was added and centrifuged at 13,000 rpm for 2 min and the supernatant were discarded. The washing was repeated twice. The silica pellets were dried in heating block at 60 °C for 10 min. 80 µl of TE buffer (10 mM Triss pH 7.8 and 1 mM EDTA) was added in each tube and mixed gently. The tubes were then centrifuged at 13,000 rpm for 5 min to collect the aqueous DNA solution.

The DNA extracted from above three methods was used for amplification of 472 bp long mtDNA cyt *b* gene fragment with universal primers (Verma and Singh 2003). This primer has been used in investigation of various crime cases related to species identification (Gupta et al. 2005, 2012). PCR reactions were carried out in 20 µl reaction volume by using Hot Start Multiplex PCR kit (Qiagen, Germany) with 4 pmol of each primer and 1 µl template DNA. The PCR conditions were: 95 °C for 15 min, followed by 35 cycles each 95 °C for 45 s, 55 °C for 40 s, and 72 °C for 1.5 min. The final extension was at 72 °C for 10 min. The PCR products were electrophoresed on 2 % agarose gel, stained with ethidium bromide (0.5 mg/ml) and visualized under U.V. transilluminator (Fig. 1).

All positive PCR products were treated with *Exonuclease-I* (*Exo-I*) and *Srimp Alkaline Phosphatase* (*SAP*) to clean the unused primers and dNTP's. Cleaned PCR products were used for sequencing with BigDye sequencing kit and 3130 ABI Genetic Analyzer (Applied Biosystem). The sequence resolved was aligned and cleaned by using Sequencher 4.7 (Gene Code Corporation, Ann Arbor, USA). The DNA sequences were used for Blast search (at http://blast.ncbi.nlm.nih.gov/) to confirm the origin of the sample and to check the purity of the DNA extraction and amplification.

Results and conclusion

The three different DNA extraction protocols yielded different PCR outcomes. PCR amplification in bone samples was negligible in PC method, low in Qiagen kit and high with Gu-HCl method (Fig. 1). Poor amplification was observed for DNA extracted from bone, antler, and feces using PC method (Fig. 1). In antlers, PCR amplifications were detected in all the three methods. For feces samples, Qiagen Stool kit and Gu-HCl method show comparable PCR amplification. The PC method was found inconsistent with the fecal samples (Fig. 1). Although 7–8 samples of each group including bone, antler, feces and tissue were used for assessment of the result, the results from only two samples of each group are shown in Fig. 1.

This study evidently indicates that two different Qiagen kits (blood/tissue and stool) were required for obtaining the higher PCR success rate from different sample types. However, Gu-HCl based DNA extraction protocol illustrated uniformly higher PCR success from all types of biological samples. The amplified PCR product yielded good DNA sequence, which showed authentic matching of the source of its origin and confirmed that the same can be applied on a variety of biological samples with uniform success rate. The uniform success rate in PCR and sequencing evidenced that Gu-HCl method is fast, low-cost, and less hazardous. This protocol was also successfully tested on antler and bone samples of Swamp deer (*Rucervus duvaucelii*) and critically endangered Sangai deer (*Rucervus eldii eldii*).

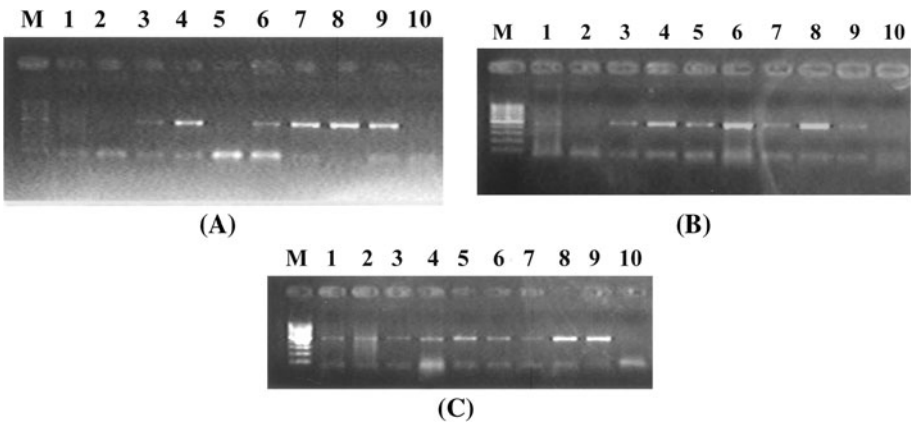


Fig. 1 PCR amplification with the DNA extracted from various sample using PC (**a**), Qiagen kit (**b**), and Gu-HCl method (**c**). *Lane M* 100 bp ladder; *Lane 1–2* Bone; *Lane 3–v4* Antler; *Lane 5–6* Feces; *Lane 7–8* Tissue; *Lane 9* and *10* are positive control and extraction negative control, respectively

Acknowledgments This work was funded by the Ministry of Environment and Forest, Govt. of India through the Grant-in-Aid to WII. We thank to the Director, Dean, and Nodal Officer (Forensic Cell) of WII for their support.

References

Gupta SK, Verma SK, Singh L (2005) Molecular insight into a wildlife crime: the case of a peafowl slaughter. Forensic Sci Int 154:214–217

Gupta SK, Thangaraj K, Singh L (2011) Identification of the source of ivory idol by DNA analysis. J Forensic Sci 56:1343–1345

Gupta SK, Thangaraj K, Singh L (2012) Molecular insight into suspected food crime cases: a study to determine the legitimacy of meat by DNA typing. Biol Biomed Rep 2:171–174

Höss M, Pääbo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Res 21:3913–3914

Rohland N, Hofreiter M (2007a) Ancient DNA extraction from bones and teeth. Nat Protoc 2:1756–1762

Rohland N, Hofreiter M (2007b) Comparison and optimization of ancient DNA extraction. Biotechniques 42:343–352

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Verma SK, Singh L (2003) Novel universal primers establish identity of enormous number of animal species for forensic application. Mol Ecol Notes 3:28–31

Wehausen JD, Ramey RR II, Epps CW (2004) Experiments in DNA extraction and PCR amplification from bighorn sheep faeces: the importance of DNA extraction method. J Hered 95:503–509