Legal Medicine 17 (2015) 539–546

Contents lists available at ScienceDirect

Legal Medicine

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

Announcement of Population Data

A study of genetic polymorphisms in mitochondrial DNA hypervariable regions I and II of the five major ethnic groups and Vedda population in Sri Lanka

Ruwandi Ranasinghe [a], Kamani H. Tennekoon [a,*], Eric H. Karunanayake [a], Maria Lembring [b], Marie Allen [b]

[a] Institute of Biochemistry, Molecular Biology & Biotechnology, University of Colombo, 90 Cumaratunga Munidasa Mawatha, Colombo 03, Sri Lanka
[b] Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Biomedical Centre, Box 815, 751 08 Uppsala, Sweden

ARTICLE INFO

Article history:
Received 14 October 2014
Received in revised form 19 May 2015
Accepted 20 May 2015
Available online 27 May 2015

Keywords:
Mitochondrial DNA
Control region
HV region I
HV region II
Haplotype
Sri Lanka

ABSTRACT

Diversity of the hypervariable regions (HV) I and II of the mitochondrial genome was studied in maternally unrelated Sri Lankans ($N = 202$) from six ethnic groups (i.e.: Sinhalese, Sri Lankan Tamil, Muslim, Malay, Indian Tamil and Vedda). DNA was extracted from blood and buccal swabs and HVI and HVII regions were PCR amplified and sequenced. Resulting sequences were aligned and edited between 16024–16365 and 73–340 regions and compared with revised Cambridge reference sequences (rCRS). One hundred and thirty-five unique haplotypes and 22 shared haplotypes were observed. A total of 145 polymorphic sites and 158 polymorphisms were observed. Hypervariable region I showed a higher polymorphic variation than hypervariable region II. Nucleotide diversities were quite low and similar for all ethnicities apart from a slightly higher value for Indian Tamils and a much lower value for the Vedda population compared to the other groups. When the total population was considered South Asian (Indian) haplogroups were predominant, but there were differences in the distribution of phylo-geographical haplogroups between ethnic groups. Sinhalese, Sri Lankan Tamil and Vedda populations had a considerable presence of West Eurasian haplogroups. About 2/3rd of the Vedda population comprised of macro-haplogroup N or its subclades R and U, whereas macro-haplogroup M was predominant in all other populations. The Vedda population clustered separately from other groups and Sri Lankan Tamils showed a closer genetic affiliation to Sinhalese than to Indian Tamils. Thus this study provides useful information for forensic analysis and anthropological studies of Sri Lankans.

Population: Sri Lanka, a Southern Asian Island, known as Ceylon until 1972, is located southwest to the Bay of Bengal in the Indian Ocean. It is inhabited by different ethnicities, having diverse religions and languages. Its population consists of the ethnicities Sinhalese (74.9%), Sri Lankan Tamils (11.2%), Muslims (Sri Lankan Moors) (09.2%), Indian Tamils (04.2%), Malays (0.2%), Burghers (0.2%), indigenous Vedda people (0.1%) and a few others including Sri Lankan Chetty and Bharatha [1]. Sri Lanka has been a vital trading point of the Indian Ocean trade network for more than 2000 years thus providing opportunities for cultural and genetic admixture. According to historical evidence, many groups have arrived in the country as invaders, immigrants and traders in pre-historical and historical eras. Linguistic studies indicate historical relationships between Sinhalese, an Indo–Aryan language, and early Maldivian languages [2]. The Dravidian language known as Tamil is spoken by Sri Lankan Tamils, Indian Tamils as well as some Muslims in Sri Lanka, especially by those living in areas dominated by Tamils.

There is archaeological evidence of prehistoric settlements dating as far back as 125,000 years before present (bp) and additional convincing support from analyses of human skeletal remains from 37,000 years bp in Sri Lanka [3]. Today the Vedda population is thought to be the remnants of the original inhabitants of Sri Lanka. Admixture of North Indians with original inhabitants is reported to have established the Sinhalese ethnic group approximately 2600 years ago, while the origin of Sri Lankan Tamils is attributed to successive invasions by South Indians beginning a few centuries later [4,5]. Furthermore, females of both North Indian and South Indian origin have arrived in the country from time to time during the last 2600 years, either as consorts for rulers (both Sinhalese and Tamil) and their Councils, or as part of the entourage accompanying Buddhist nuns. Muslims from the

* Corresponding author. Tel.: +94 112552534; fax: +94 112552529.
E-mail addresses: ruwandiranasinghe@gmail.com (R. Ranasinghe), kamani@ibmbb.cmb.ac.lk (K.H. Tennekoon), erick@ibmbb.cmb.ac.lk (E.H. Karunanayake), maria.lembring@mail.com (M. Lembring), marie.allen@igp.uu.se (M. Allen).

http://dx.doi.org/10.1016/j.legalmed.2015.05.007
1344-6223/

Arabian Peninsula arrived in the country between 7th to 8th centuries, mostly as traders, whereas Malays from South East Asia arrived during the Dutch occupation of Sri Lanka in the 17th and 18th centuries. Indian Tamils were brought to Sri Lanka mainly from Tamil Nadu by the British during the 19th Century. Burghers are descendants of Europeans who invaded and occupied Sri Lanka from 16th to the 20th Century [6].

Genetic variations and gene flow of Sri Lankans are still being investigated. Studies based on classical and autosomal markers give somewhat contradictory evidence on population relationship between ethnic groups. The major ethnic group Sinhalese, were found to be genetically closer to Bengalis than to Tamils, Gujaratis, or Punjabis by some [7], while others reported that Sinhalese and Sri Lankan Tamils are closely related [8]. A study on thrombophilic polymorphisms [9] and another on genetic profiling of 11 autosomal microsatellite markers [10] showed that Sinhalese are closely related to both Sri Lankan Tamils and Muslims, while a study on candidate genes for pre-eclampsia indicated significant genetic differences between Sinhalese and Muslims [11].

A recent phylogenetic study on mitochondrial DNA hypervariable segments HVI and part of HVII showed that the Vedda people are genetically distinct from other major ethnic groups (Sinhalese, Sri Lankan Tamils and Indian Tamils) in Sri Lanka [12]. Another study on alpha-2-HS-glycoprotein allele frequency has shown that Veddas and Sinhalese are more biologically related to each other than to other ethnic groups (Thai and Paraguayan) from Asia [13]. A previous study on Vedda people and four other Asian populations (Japanese, Ainu of Northern Japan, Korean, Negrito of the Philippines) also showed that the Vedda population clusters separately from the others when an UPGMA tree was constructed based on genetic distances obtained following restriction digestion with 13 enzymes [14] suggesting a longer isolation of Vedda population from other older Asian inhabitants. Chaubey [15] has emphasized the value of studying prehistoric settlements of Sri Lankans and their relation with neighboring countries as this might yield an insight into initial settlers who came from southern migration route to the South Indian subcontinent.

Human migration has been studied using the maternally inherited mitochondrial genome. The major mtDNA haplogroups of modern human are macro-haplogroups L, M and N. Evolutionary studies show that mtDNA variants stemmed from Africa about 200,000 years ago [16]. Macro-haplogroup L considered the African-specific haplogroup is limited to the sub-Saharan African region and this is further divided into distinctive subclades, L0 to L6 [17]. L3 is considered as the out of Africa haplogroup and macro-haplogroups M and N are sibling haplogroups derived from the root L3. The highest global distribution of haplogroups M and N has been traced to the South Asian region, especially the Indian subcontinent [18]. The macro haplogroup M is the most common haplogroup present in Asia [19]. More than 60% of Indians have their maternal roots in Indian- specific branches of haplogroup M [20]. In contrast, haplogroup N exists in almost all European, Oceanian, and many Asian and Amerindian populations [21,22].

Analysis of mitochondrial DNA of different ethnic groups in Sri Lanka, haplogroup determination and statistical interpretation will provide genetic evidence for the relationship of maternal lineage of ethnic groups in contemporary Sri Lanka. However, there has been only one study on mtDNA on Sri Lankan ethnic groups to date, and this was limited to Sinhalese, Tamil and Vedda populations [12]. The objective of the present study was to analyze the sequence data of the two hypervariable regions of human mtDNA from six ethnic groups in Sri Lanka including the Vedda population in order to ascertain the matrilineal genetic contribution of these groups to the present-day gene pool.

Blood or buccal samples were collected from 202 healthy, randomly selected, maternally unrelated volunteers: Sinhalese ($N = 60$), Sri Lankan Tamil ($N = 30$), Indian Tamil ($N = 22$), Muslim ($N = 30$), Malay ($N = 30$) and Vedda ($N = 30$). Malays were selected as they originated from South East Asia, Indian Tamils in view of their recent migration from Southern India and the other four groups in view of the long history of settlement in Sri Lanka. Sinhalese, Muslims and Malays were mostly from the Western Province. Sri Lankan Tamils were either resident in the Northern Province or had recently migrated to the Western Province from the Northern Province, while all Indian Tamils were from the Central Province (Fig. 1). Vedda individuals were from Dambana and Henanigala in the Uva and Eastern Provinces respectively. Ethical approval was obtained from the Research Ethics and Higher Degrees Committee of the Institution and verbal informed consent was obtained from the study participant.

DNA extraction and quantification: DNA was obtained from blood samples using Wizard® Genomic DNA Purification Kit (Promega Corporation, Madison, USA) according to the manufacture's protocols and from Buccal samples as described by Saab and others [23]. After completing the extraction each sample was loaded into 0.8% agarose gel (with 0.01% of ethidium bromide) and run constantly for about 2–3 h at 40 V in 1X TAE buffer with λ concentration markers (20, 50, 100 and 200 ng/μl). After electrophoresis, products were visualized using gel documentation system (QUANTUM ST4® imaging systems, Viber Lourmat SAS, France) to estimate the quantity of DNA.

PCR amplification and sequencing: PCR amplifications were performed using rTaq DNA polymerase (GE Healthcare, Buckinghamshire, UK) to amplify the 440 bp fragment with 15971F/16410R primers for the HVI region and 415 bp fragment with 15F/429R primers for the HVII region [24]. PCR amplification was carried out in a reaction mixture of 25 μl, containing 4 mM $MgCl_2$, 1× PCR buffer, 0.2 mM each dNTP, 4 pmol of each primer, 100–200 ng of total DNA and 1 U of rTaq polymerase. PCR cycling of the target fragments was carried out in a master cycler (Eppendorf-Mastercycler gradient, USA) under the following conditions: 95 °C for 7 min followed by 38 cycles of 94 °C for 20 s, 56 °C for 10 s, and 72 °C for 30 s, final extension at 72 °C for 10 min. DNA from buccal swabs was amplified by using PCR conditions described previously [25] with minor modifications (2.5 mM $MgCl_2$, 5–100 ng of DNA and 1.3 U of rTaq polymerase) and cycle parameters used were the same as for blood samples. Along with every set of PCR, no template reactions were amplified as negative controls.

Sequencing: Amplified products were purified by using illustra GFX™ PCR DNA and Gel Band purification Kit (GE Healthcare) and the concentration of purified products was estimated on 1% agarose gel. PCR products were sequenced with the PCR primers (described above), using DYEnamic™ ET Dye Terminator Cycle sequencing Kit for MegaBACE™ (GE Healthcare), according to the manufacture's protocol. After thermal cycling, the reaction products were ethanol precipitated and electrophoresed on MegaBACE1000 automated DNA sequencer (GE Healthcare). Duplicate PCR and sequencing reactions were performed for each sample and both forward and reverse strands were sequenced to reduce ambiguities.

Analysis of data: Nucleotide positions were numbered according to the revised Cambridge Reference sequence [26,27] and edited, aligned, and compared both manually and electronically between nucleotide positions 16024–16365 for HVI and 73–340 for HVII, using BioEditv7.0.9 [28] and Sequencher® 4.10.1-Build 5829 (Gene Codes Corporation, Ann Arbor, MI, USA) software respectively by two independent individuals. Single nucleotide polymorphisms were analyzed and reported by using Mutation