

RESEARCH ARTICLE

# Sri Lankan maternal ancestry reveals early migrations from Africa along the Indian Ocean

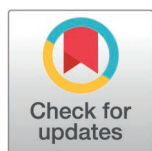
Anjana Welikala<sup>1</sup>\*, Shailesh Desai<sup>2</sup>\*, Amali Fernando<sup>1</sup>, Joanne Kotelawala<sup>1</sup>, Lakshika Jayasekara<sup>1</sup>, Prajval Pratap Singh<sup>2</sup>, Anoma Jayasoma<sup>1</sup>, Jose Antonio Urban Aragon<sup>3</sup>, David Witonsky<sup>3</sup>, Niraj Rai<sup>4</sup>, Kumaraswamy Thangaraj<sup>4</sup>, Maanasa Raghavan<sup>3</sup>, Gamini Adikari<sup>5</sup>, Kamani H. Tennekoon<sup>1</sup>, Gyaneshwer Chaubey<sup>2\*</sup>, Ruwandi Ranasinghe<sup>1\*</sup>

**1** Institute of Biochemistry, Molecular Biology and Biotechnology, University of Colombo, Colombo, Sri Lanka, **2** Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi, India, **3** Department of Human Genetics, University of Chicago, Cummings Life Science Center, Chicago, Illinois, United States of America, **4** CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, **5** Postgraduate Institute of Archaeology, University of Kelaniya, Colombo, Sri Lanka

\* These authors contributed equally to this work.

✉ Current address: Birbal Sahni Institute of Palaeosciences, Hasanganj, Lucknow, India.

\* [gyaneshwer.chaubey@bhu.ac.in](mailto:gyaneshwer.chaubey@bhu.ac.in) (GC); [anjana@ibmbb.cmb.ac.lk](mailto:anjana@ibmbb.cmb.ac.lk) (AW); [ruwa@ibmbb.cmb.ac.lk](mailto:ruwa@ibmbb.cmb.ac.lk) (RR)



## OPEN ACCESS

**Citation:** Welikala A, Desai S, Fernando A, Kotelawala J, Jayasekara L, Singh PP, et al. (2026) Sri Lankan maternal ancestry reveals early migrations from Africa along the Indian Ocean. PLoS One 21(5): e0350045. <https://doi.org/10.1371/journal.pone.0350045>

**Editor:** Jun Gojobori, SOKENDAI (The Graduate University for Advanced Studies), JAPAN

**Received:** July 24, 2025

**Accepted:** May 6, 2026

**Published:** May 26, 2026

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data availability statement:** All relevant data are within the manuscript and its [Supporting Information](#) files.

**Funding:** The authors received no specific funding for this work.

## Abstract

The African origin of anatomically modern humans is widely accepted. However, there is ongoing debate about the route they took and whether the early expansion into Oceania was through South Asia. Despite Sri Lanka being an island South Asian nation with the earliest known human fossils in South Asia and strategically located along a putative 'southern route', a comprehensive examination of its temporal settlement using high-resolution complete mitochondrial DNA analysis has never been conducted. To address this gap, 139 mitogenomes were sequenced in this study from the Sinhalese, Sri Lankan Tamil, and Vedda populations in Sri Lanka and integrated with 247 previously published global mitogenomes, resulting in the largest mitogenome dataset analyzed thus far. Phylogeographic analyses revealed four distinct settlement phases in Sri Lanka, with the earliest phase overlapping with the initial entry of modern humans into South Asia, thus supporting the southern dispersal route. The introduction of West Eurasian lineages into Sri Lanka was mediated via India. A significant decline in effective population size was observed across all studied populations, reflecting the demographic history of the island. Findings from the present study provide valuable insights into the long-standing debate on the southern and inland migration routes out of Africa and subsequent migrations from across Eurasia, thus highlighting the complex settlement patterns of Sri Lanka and broader Asia.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Sri Lanka, an island nation surrounded by the Indian Ocean, stands as a mosaic of cultural diversity within the rich socio-cultural tapestry of South Asia. The country is positioned in the southeastern expanse of the Indian subcontinent between 5°55' and 9°51' N, encompassing over ~65,000 km<sup>2</sup> [1]. The country's topography can be divided into three distinct regions based on elevation: the central highlands (~2,500 m above sea level), plains (>200 m above sea level), and the coastal belt (>30 m above sea level). Each zone possesses its unique micro-climate, which is significantly shaped by factors such as temperature, wind patterns, ocean currents, rainfall, and vegetation [2]. Both the geography and topology of Sri Lanka have influenced human migrations into and within the country over time.

The 2024 Census and statistics [3] reported that the country is inhabited by over 22 million people from diverse ethnic groups. Sinhalese represents the majority, comprising of approximately 75% of the total population, followed by Sri Lankan Tamil (SLT) (~11%), Sri Lankan Moor (~9%), Indian Tamils (~4%), and minority populations such as Malays, Burghers, and Veddhas (~1%). The country has held a strategic position at the crossroads of Indian Ocean trade routes, witnessing population movements since prehistoric and historic times [4–7]. Despite being an island today, Sri Lanka was impermanently coupled to the southern tip of the Indian subcontinent through a transient land bridge during epochs of lower sea levels [8]. This connection facilitated the rapid movement of human populations since at least the Paleolithic period [9–10].

The present-day biocultural landscape of Sri Lanka is a result of an admixture of diverse populations that have migrated to the island at various times. The earliest traces of human settlements in Sri Lanka have been dated to approximately 48,000 years before the present (YBP) [11]. These ancient inhabitants were the Mesolithic hunter-gatherers, who established their homes in tropical caves and arid open-air sites until ~7,000 BP [9,12–14]. Both archaeological and anthropological evidence demonstrate that the present-day Vedda peoples of Sri Lanka are descendants of these Mesolithic hunter-gatherers who later admixed with the Iron Age populations that migrated to Sri Lanka during the proto and early historic periods [9,15–18]. Such early evidence of anatomically modern humans in Sri Lanka, coupled to its coastal location highlights Sri Lanka's critical role in the early peopling of South Asia as well as the proposed southern route out of Africa.

The chronology of human settlements during the proto, early, and later historic periods is primarily derived from the ancient chronicles and stone inscriptions [19–21]. Embedded in cultural beliefs, the Sinhalese who speak an Indo-Aryan language (Sinhala) are said to trace their origin to the legendary Prince Vijaya and his followers, who arrived from India during the proto-historic period (543 BCE). In contrast, the Dravidian-speaking SLTs originated in southern India and established connections with the island through marriages, invasions, trading, and other interactions during the historical periods [19]. Notably, the Veddhas, who speak a linguistic isolate, stand as the sole Indigenous/Adivasi population in present-day Sri Lanka, representing the enduring generations of stone-age people [9,15–18,22].

Over millennia, Sri Lanka has witnessed extensive population movements between its shores and neighboring regions, shaping a dynamic tapestry of biological and cultural exchanges. While most of these migrations were predominantly male-mediated, there have been notable, though comparatively rare instances of female-mediated gene flow, as documented in historical accounts [19,21]. Despite these known interactions, the precise impact of such historical events on the genetic makeup of the Sinhalese, SLT, and Vedda populations remains a topic of discussion, necessitating further investigation. Several genetic studies have been conducted on different ethnic groups in Sri Lanka using short tandem repeats (STRs) [23,24], mitochondrial DNA (mtDNA); hyper variable segment (HVS) I, HVSII, and dinucleotide repeat variations [25–27], X-chromosomal STRs [28], genotyping data [29], and whole genomes [30]. Previous studies based on autosomal STR and mitochondrial HVS regions revealed that the Vedda population is genetically distinct from other ethnic groups in Sri Lanka [24–26]. Notably, they exhibit a high prevalence of Eurasian mitochondrial haplogroups (U and R) compared to South Asian haplogroups. Conversely, both Sinhalese and SLT displayed a mix of South Asian and Eurasian haplogroups with a higher prevalence of Indian haplogroups. Furthermore, recent studies utilizing genome-wide data have reported substantial genetic similarity between the Sinhalese and SLTs, despite their distinct linguistic and cultural backgrounds. This gene flow appears to have originated from both North and South Indian populations [29,30]. In contrast, the Vedda people have maintained a degree of genetic isolation and low effective population sizes, reflected by their high genetic drift, with high genetic similarity to Indian tribal populations likely mediated through admixture with migrating groups from India, such as the Sinhalese and SLT [22,30]. However, despite the potential impact of sex-biased migrations on the genetic structure of Sri Lankans, a comprehensive investigation utilizing high-resolution complete mtDNA has not yet been conducted to unravel the maternal genetic lineages and the phylogeography of the Sinhalese, SLTs, and Vedda populations.

To address this gap, we sequenced 139 mitogenomes from Sinhalese, SLT, and Vedda populations in this study, comparing them with 247 previously published mitogenomes [31–33]. This comprehensive dataset was analyzed to address two primary questions. First, given that Sri Lanka is home to the oldest evidence of anatomically modern humans in South Asia and is located at the crossroads of the proposed “southern route migration” out of Africa, can this data contribute to our understanding of the ongoing debate between the southern and inland migration routes? Second, what insights do the maternal ancestries in Sri Lanka provide on the timeline of migrations from East Eurasia, West Eurasia, and mainland South Asia?

## Materials and methods

### Sample collection and DNA extraction

Blood samples were collected from 139 healthy, maternally unrelated individuals [Sinhalese (N=96), SLTs (N=6), and Vedda (N=37)] after obtaining both verbal and written informed consent between 15<sup>th</sup> March 2021 and 21<sup>st</sup> November 2022. The ethical approval for the study was granted by the Ethics Review Committee of the Faculty of Medicine, University of Colombo (EC-17–147). In addition, A total of 103 complete mitochondrial genome sequences of SLTs were retrieved from the 1000 Genomes Project and incorporated into the analysis, resulting in a total SLT sample size of 109 [32].

The blood DNA extraction was carried out, either following established protocols [26,34] or using commercially available DNA extraction kits [Dneasy Blood & Tissue DNA extraction kit (Catalogue no: ID: 51304 Qiagen, Hilden, Germany), QIAmp DNA Blood Mini kit, DNA Investigation kit (Catalogue no: ID: 51104 Qiagen, Hilden, Germany)]. The quality and quantity of the extracted DNA were assessed using a Nano-spectrophotometer (Shimadzu BioSpec-Nano™ spectrophotometer, Japan).

### Mitogenome sequencing and haplogroup assignment

Mitogenome sequencing was performed with two distinct methodologies: direct sequencing and next-generation sequencing (NGS). Direct sequencing was applied to 52 samples, involving PCR amplification of mitogenomes using 24 specific

primers [35,36]. PCR amplification was conducted using the 96-well plate protocol, outlined in Welikala et al [27]. The PCR products were subjected to Sanger sequencing using the BigDye® Terminator v3.1 cycle sequencing kit (Catalog number 4337455: Thermo Fisher Scientific Inc., USA,) and ABI 3500Dx Genetic Analyzer (Thermo Fisher Scientific Inc., USA). Mitogenomes alignment and analysis were performed against the revised Cambridge Reference Sequence (rCRS; GenBank Accession Number NC\_012920.01) [37] using Unipro UGENE version 35.0 [37]. Mutations and their positions were then confirmed using the MITOMAP database [38].

The remaining 87 samples were subjected to NGS, utilizing the Illumina Novaseq 6000 platform at the University of Chicago, USA, and at a commercial NGS service facility (Genotypic Technology, Bangalore, India). All reads were mapped to the rCRS reference using bwa v0.7.17 and the parameters noted in previous studies [39–40]. The sample details, their haplogroup classification, and sequence polymorphisms of each sample are listed in the [S1 Table](#).

Haplogroup assignment was performed based on observed variants within 242 complete mitochondrial genomes of the three study populations [Sinhalese (N=96), SLTs (N=109), Vedda (N=37)] using HaploGrep version 3 [41], PhyloTree build 17 [31], and updated M, N, and U7 trees [33,42]. Manual verification was conducted for all mutations and haplogroups.

### Genetic diversity analysis and, phylogenetic reconstruction

Molecular diversity indices for the three study populations were computed using DnaSP version 6 [43] and Arlequin 3.5.2.2 [44]. Haplogroup frequencies were calculated using the direct counting method. Phylogenetic trees for each haplogroup were constructed using MtPhyl [45] and manually edited where necessary. The constructed trees were incorporated with additional sequences (N=144) from Indian populations as well as other South Asian and worldwide populations ([S2 Table](#)). We employed the maximum parsimony approach, confirming the nomenclature of the PhyloTree build 17 [31] and reduced-median-network (RMN) approach using Network 10 (Fluxus Technology Ltd.) ([S1 Fig](#)). Novel sub-haplogroups were identified based on additional mutations observed in the individuals. Rho ( $\rho$ ) statistics [46] and the Bayesian method [47] were used to estimate the temporal trajectories of these haplogroups. Rho (with corresponding standard errors) and age (with 95% lower and upper bounds) estimations were calculated using Network 10 (Fluxus Technology Ltd.) and 'genetic distance value conversion calculator' ( $1.67 \times 10^{-8}$  substitution per site per year) [48], respectively. These estimations involved all substitutions, excluding rCRS-defining mutations [49] and the variants mentioned in the PhyloTree build 17 [31].

### Bayesian skyline plot (BSP) and Bayesian tree

To reconstruct the maternal population history and effective population size fluctuation of all three Sri Lankan groups, Bayesian skyline analysis was performed on the mitogenomes using BEAST 2.7.5 version [47]. The data were analyzed for the combined cohort as well as separately for each population to investigate population fluctuations within each group. To select the mitochondrial evolutionary model, we employed Jmodeltest 2 [50]. Based on the lowest BIC value, an evolutionary model was chosen for each population: Combined cohort (N=242) (HKY+I+G), Sinhalese (N=96) (HKY+I+G), SLTs (N=109) (GTR+I+G), Vedda (N=37) (HKY+I). An MCMC chain was run for 100 million iterations, and the generated log files were examined with tracer. Effective sample size (ESS) for each parameter was ensured to be >200. The final tree was generated using Treeannotator with the first 30% runs set as burn-in and visualized in Figtree and iTOL [51]. For BSP, table values were extracted using tracer and R packages [52], and plots were generated for the three populations.

### Principal component analysis

In order to visualize the genetic relation between populations via genetic affinity, we performed the principal component analysis (PCA) using haplogroup frequency of different groups ([S3 Table](#)), the main representative populations were taken from the Genome Asia database, and our data were subsequently adapted to original data. Using R packages ggplot2, the final PCA was plotted ([https://github.com/Shailshdesai76/script/blob/main/PCA\\_MT\\_haplogroups.R](https://github.com/Shailshdesai76/script/blob/main/PCA_MT_haplogroups.R)).

## Results

### Haplogroup distributions and prevalence in the Sri Lankan study populations

The conservation of the maternal gene pool in the Sinhalese, SLTs, and Vedda populations of Sri Lanka over millennia was examined in this study, in light of phylogenetic evidence from neighboring Indian populations [53].

High maternal diversity was observed across the studied populations, represented by major haplogroups M, R, U, H, N, A, I, and J. The haplogroup M was the most frequent haplogroup (0.508), followed by R (0.231), U (0.173), H (0.066), and N (0.012). Haplogroups A, I, and J were the least common, and each exhibited a frequency of 0.004. At a macro level, all three populations predominantly carry the Indian-specific M and R haplogroups, reflecting shared maternal genetic origins. Our fine-scale haplogroup analysis revealed intriguing relationships among populations within Sri Lanka, as well as with different Indian populations (S1 File).

In the Sinhalese population, major haplogroup M was the most prevalent, comprising 0.635 of the population. This was followed by haplogroups R (0.177), U (0.146), H (0.031), and I (0.011). Similarly, in the SLTs, haplogroup M was the most common (0.486), with lower frequencies of R (0.211), U (0.138), H (0.119), N (0.028), A (0.009), and J (0.009). Compared to Sinhalese, SLTs exhibited lower percentages of M and U haplogroups and higher percentages of R and H haplogroups. The high prevalence of the M haplogroup in both Sinhalese and SLTs resembles the widespread distribution of this haplogroup in Indian mainland populations [53–55]. In contrast, the most frequent haplogroup in the Vedda population was R at 0.432, followed by U (0.324) and M (0.243) (Table 1).

At a finer scale, a total of 103 mitochondrial sub-lineages were identified among the 242 Sinhalese, SLTs, and Vedda individuals under study. These haplogroups belong to the major haplogroups A1, H2, H6, H13, HV14, I, J1, M, M2, M5, M3, M4\*67, M4, M6, M18'38, M18, M30, M33, M34, M35, M36, M37, M38, M40, M41, M42, M44, M45, M52, M53, M65, M66, M81, N5, N21, R, R5, R6, R7, R8, R30, R31, U1, U3, U5, U7. Among them, 32 major haplogroups were observed in the Sinhalese cohort (N=96), 37 in SLTs (N=109), and 9 in Veddass (N=37).

Among the total mitochondrial haplogroups, 18 major haplogroups (M, M3, M4, M6, M18'38, M30, M33, M36, M52, M53, M44, M66, R6, R7, R8, R31, H13, U2) are shared between the Sinhalese and SLTs. Haplogroup U3 was common to both the Sinhalese and Vedda, and seven major haplogroups (M2, M5, M65, M35, R5, R30, U7) were shared across all three populations (Fig 1). Our haplogroup frequency-based PCA also revealed that all three Sri Lankan populations cluster closely together compared to other populations (S2 Fig).

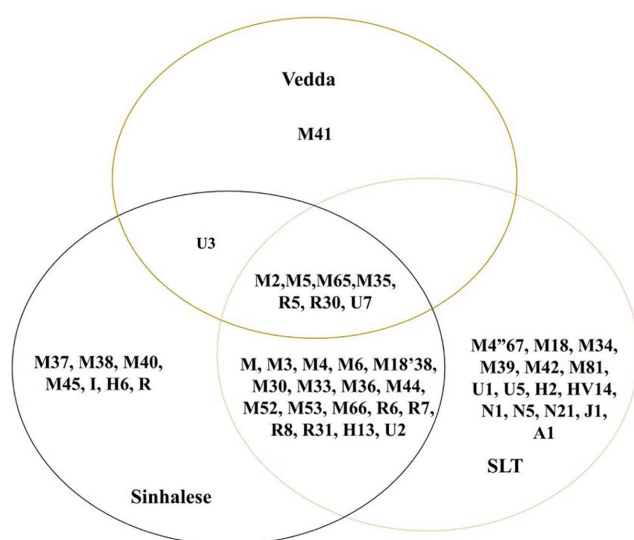
### Maternal molecular diversity in the Sri Lankan study populations

SLTs exhibited the highest average number of nucleotide differences (k) and nucleotide diversity ( $\pi$ ) among all three study populations. The Sinhalese population displayed intermediate levels while the Vedda population exhibited the lowest values (Table 2).

**Table 1. Frequencies of major haplogroups in the three study populations.**

| Major Haplogroups | Combined (N=242) | Sinhalese (N=96) | SLTs (N=109) | Vedda (N=37) |
|-------------------|------------------|------------------|--------------|--------------|
| M                 | 0.508            | 0.635            | 0.486        | 0.243        |
| R                 | 0.231            | 0.177            | 0.211        | 0.432        |
| U                 | 0.173            | 0.146            | 0.138        | 0.324        |
| H                 | 0.066            | 0.031            | 0.119        | –            |
| N                 | 0.012            | –                | 0.028        | –            |
| A                 | 0.004            | –                | 0.009        | –            |
| I                 | 0.004            | 0.011            | –            | –            |
| J                 | 0.004            | –                | 0.009        | –            |

<https://doi.org/10.1371/journal.pone.0350045.t001>



**Fig 1. Mitochondrial haplogroup sharing between the three study populations.**

<https://doi.org/10.1371/journal.pone.0350045.g001>

**Table 2. Genetic diversity indices in three study populations.**

| Population | N   | Genetic diversities         |                  |                  |
|------------|-----|-----------------------------|------------------|------------------|
|            |     | The sum of Square Frequency | K (SD)           | π (SD)           |
| Sinhalese  | 96  | 0.0104                      | 36.998 (±16.228) | 0.0022 (±0.0011) |
| SLTs       | 109 | 0.0092                      | 38.808 (±16.993) | 0.0023 (0.0011)  |
| Vedda      | 37  | 0.0270                      | 31.661 (±14.135) | 0.0019 (±0.0009) |

N- number of samples; SD- standard deviation; k- average number of nucleotide differences; π- nucleotide diversity.

<https://doi.org/10.1371/journal.pone.0350045.t002>

AMOVA indicates that 4.93% of the genetic variation is due to differences among the three study populations, while 95.07% of the variation occurs within populations. The fixation index ( $F_{ST}$ ) of 0.04928 reflects low genetic differentiation, suggesting that the majority of the genetic diversity is shared within populations, with only a small fraction attributable to inter-population differences (Table 3). These results support recent shared genetic ancestry and historical gene flow between these populations.

Additionally, the pairwise  $F_{ST}$  analysis between the populations reveals that the Sinhalese and SLT populations exhibit very low genetic differentiation, suggesting they are genetically similar, likely due to high gene flow or recent common ancestry. In contrast, the Sinhalese-Vedda and SLTs-Vedda comparisons exhibited a moderate level of genetic distinctiveness between the respective populations (Table 4).

### Novel sub-haplogroup identification

Fifteen previously unreported haplogroups were identified in the data generated in this study (Table 5). They were classified into distinct lineages according to the nomenclature of Phylotree Build 17 [31]. These new lineages were designated only when they were observed in at least two individuals [56]. The identification of these novel sub-haplogroups presents significant potential to refine the current phylogenetic framework. To ensure clarity, these novel haplogroups were highlighted in the phylogenetic trees constructed using the maximum parsimony approach (S1 Fig). These novel

**Table 3. AMOVA results of the total study populations (N=242).**

| Source of variation | Degrees of freedom (d. f.) | Sum of Squares | Variance Components | Percentage of variation |
|---------------------|----------------------------|----------------|---------------------|-------------------------|
| Among populations   | 2                          | 179.403        | 0.95556 Va          | 4.93                    |
| Within populations  | 239                        | 4405.527       | 18.43317 Vb         | 95.07                   |
| Total               | 241                        | 4584.930       | 19.38873            |                         |

Fixation Index ( $F_{ST}$ ): 0.04928; p: 0.0000.

<https://doi.org/10.1371/journal.pone.0350045.t003>

**Table 4. Pairwise ( $F_{ST}$ ) distances between the study populations.**

|           | Sinhalese | SLTs    | Vedda |
|-----------|-----------|---------|-------|
| Sinhalese | —         |         |       |
| SLTs      | 0.01735   | —       |       |
| Vedda     | 0.10307   | 0.07802 | —     |

p < 0.05; Sinhalese (N = 96); SLTs (N = 109), Vedda (N = 37).

<https://doi.org/10.1371/journal.pone.0350045.t004>

sub-haplogroups likely represent recent divergences, potentially tracing back to the last ~10,000 years based on Rho estimated ages, with some branches possibly differentiating within the island as recently as the last few thousand years. In addition, a few individuals were observed with a unique set of variants that appeared only once and were thus considered private mutations.

## Demographic expansion of maternal lineages in Sri Lanka

Tajima's D and Fu's  $F_s$  values exhibited significantly negative values for all three study populations, with p-values less than 0.05, indicating statistical significance. However, the Vedda population exhibited slightly lower values of these statistics compared to Sinhalese and SLTs, with a non-significant p value (0.073) for Tajima's D. Bayesian Skyline analysis revealed that the maternal lineage expansion of Sinhalese and SLT occurred approximately 45,000 YBP (Fig 2), aligning with the expansion timeline observed in Indian populations [33,55]. In contrast, the Vedda population's expansion occurred later, around ~35,000 YBP (Fig 2). Moreover, the effective population size ( $N_e$ ) of the Vedda population has remained low both before and after their demographic expansion compared to the Sinhalese and SLTs.

The demographic expansion of the Sinhalese and SLTs occurred around 45 kilo years ago (KYA), whereas the Veddas experienced expansion around 35 KYA. A decline in  $N_e$  in the Sinhalese and SLTs began after ~5 KYA, while in the Veddas the decline in  $N_e$  started much earlier, around ~10 KYA.

## Discussion

While there is consensus on the African origins of anatomically modern humans, there is considerable debate about the number of migration events [54], the timing of these events, and the routes our ancestors took as they travelled across Eurasia to reach Oceania [57–61]. In particular, the discussion surrounding the initial migration routes of humans from Africa, whether via inland routes or coastal avenues, remains contentious [55,58,59] in large part due to rising ocean levels, limiting access to archaeological evidence along the Indian Ocean coastline. To examine the southern versus inland route models taken by anatomically modern humans after they left Africa, we leveraged a large dataset of mtDNA from an underrepresented yet crucial region of South Asia, namely Sri Lanka.

We undertook haplogroup distribution and phylogenetic analyses to shed light on temporal shifts in maternal ancestries in Sri Lanka and how these relate to both early and more recent human dispersals into the region.

**Table 5. Identification of novel sub-haplogroups based on common variants in samples.**

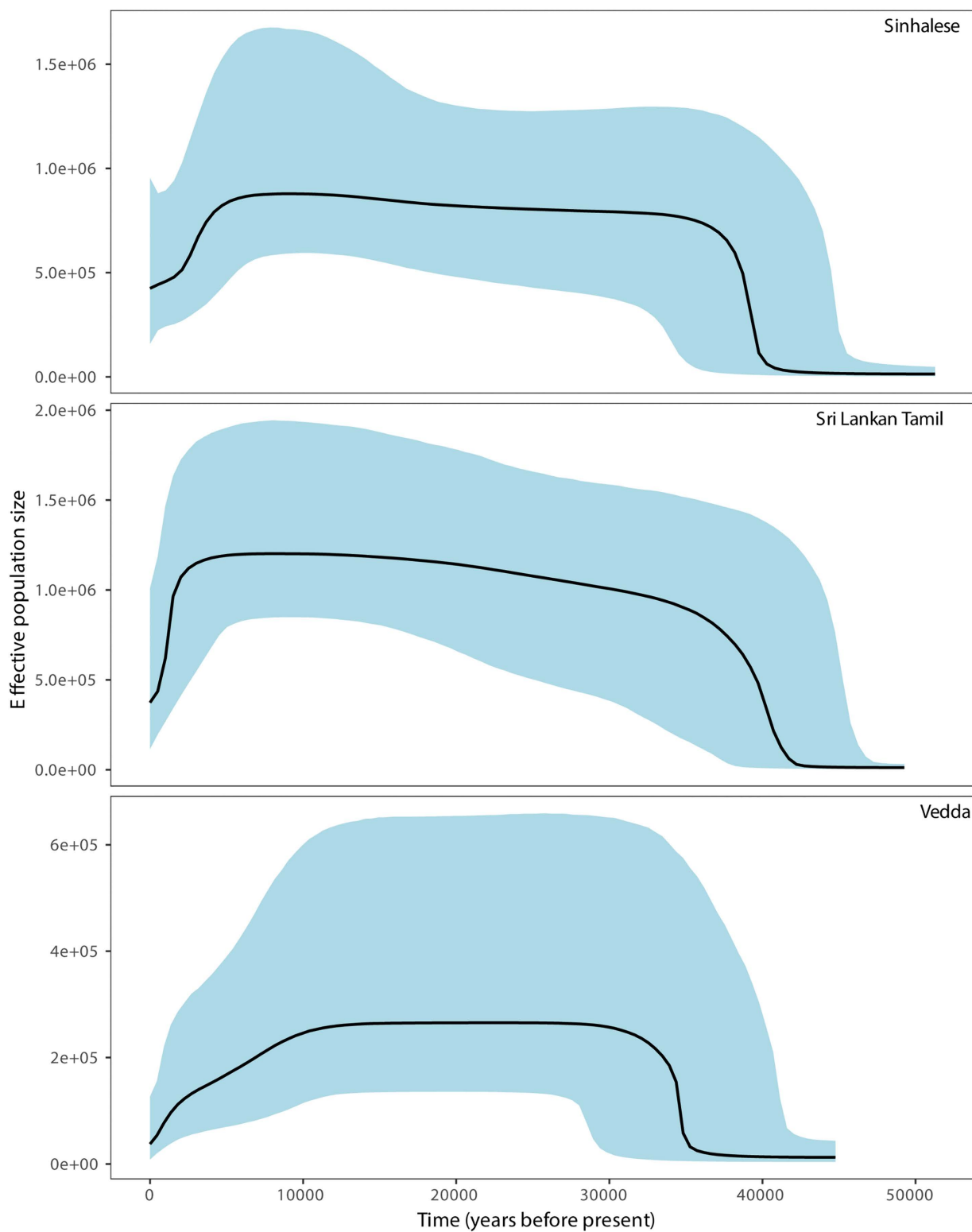
| Sample name                                                          | Haplogroup given from Haplogrep | Newly assigned Haplogroup | Additional mutations defining novel sub-haplogroups                                                                          |
|----------------------------------------------------------------------|---------------------------------|---------------------------|------------------------------------------------------------------------------------------------------------------------------|
| HG03849, HG03692, HG03995                                            | H13a2a                          | H13a2a2a                  | T152C, A4021T                                                                                                                |
| S11, mts45                                                           | M2a1                            | M2a1d                     | A10398G, G15301A, C16223T                                                                                                    |
| HG04042, HG03951, HG03857                                            | M2a1                            | M2a1d1                    | C10400T                                                                                                                      |
| HG03991, HG03955                                                     | M5a'b*                          | M5a1b1                    | T7999C, A15902G                                                                                                              |
| HG03990, HG03846                                                     | M30f                            | M30f1                     | A4487G, G8557A                                                                                                               |
| HG04075, HG03896, seq21                                              | M36                             | M36d2                     | T850C, C2380T, G3834A, A3865G, A4638G, A5843G, T6320C, G8065A, A11065G, C12302T, C12348T, T12732C, C14881T, C15493T, C16193T |
| seq28, CS06                                                          | M37e                            | M37e3                     | G35A, G36A, T146C, T634C, T2626C, T14302C, C16184T, C16185T                                                                  |
| VE14, VE03                                                           | M41                             | M41d                      | C500T, C4775G, A10398G, G15172A, G15301A, T16136C, T16172C, C16223T                                                          |
| HG03679, HG03680, V3, VE02, VE05, VE06, VE07, VE09, VE12, VE15, VE17 | R30b2a                          | R30b2a1                   | A16497G                                                                                                                      |
| S9, seq55, mts31                                                     | U2a1a                           | U2a1a1a1                  | T8227C                                                                                                                       |
| S14, VE01                                                            | U2c1a                           | U2c1a1                    | G143A, C4730T, G8020A, A13105G, C16278T                                                                                      |
| HG04099, HG03691                                                     | U2c1a                           | U2c1a1a                   | A13966G                                                                                                                      |
| V64, VE04, VE08, VE10, V4, V2, V25, VE13                             | U7a2                            | U7a2b                     | G16129A                                                                                                                      |
| CT20, HG03681, HG03711                                               | U7a2                            | U7a2c                     | G15355A                                                                                                                      |
| seq51, HG03696                                                       | U7a3a                           | U7a3a1a2a                 | T824C, T2863C, T6620C, C7028T, C16069T, G16274A, A16318C                                                                     |

<https://doi.org/10.1371/journal.pone.0350045.t005>

First, our research confirmed the presence of several previously documented haplogroups within South Asia, specifically haplogroups M2, M6, M18, M35, M36, M38, R6, R7, R30, and U2 (Fig 3 and S3 Fig). The expansion time of these haplogroups in Sri Lanka largely overlaps with the entry of modern humans in South Asia (S1 and S4 Figs; S4 Table). In addition, we refined the South Asian mtDNA tree and uncovered several sub-haplogroups nested within 14 major haplogroups present in Sri Lanka (S1 Fig). The presence of these deep-rooted haplogroups is exclusive to South Asia, suggesting minimal or no direct gene flow from other geographic regions into Sri Lanka during the initial peopling of Sri Lanka.

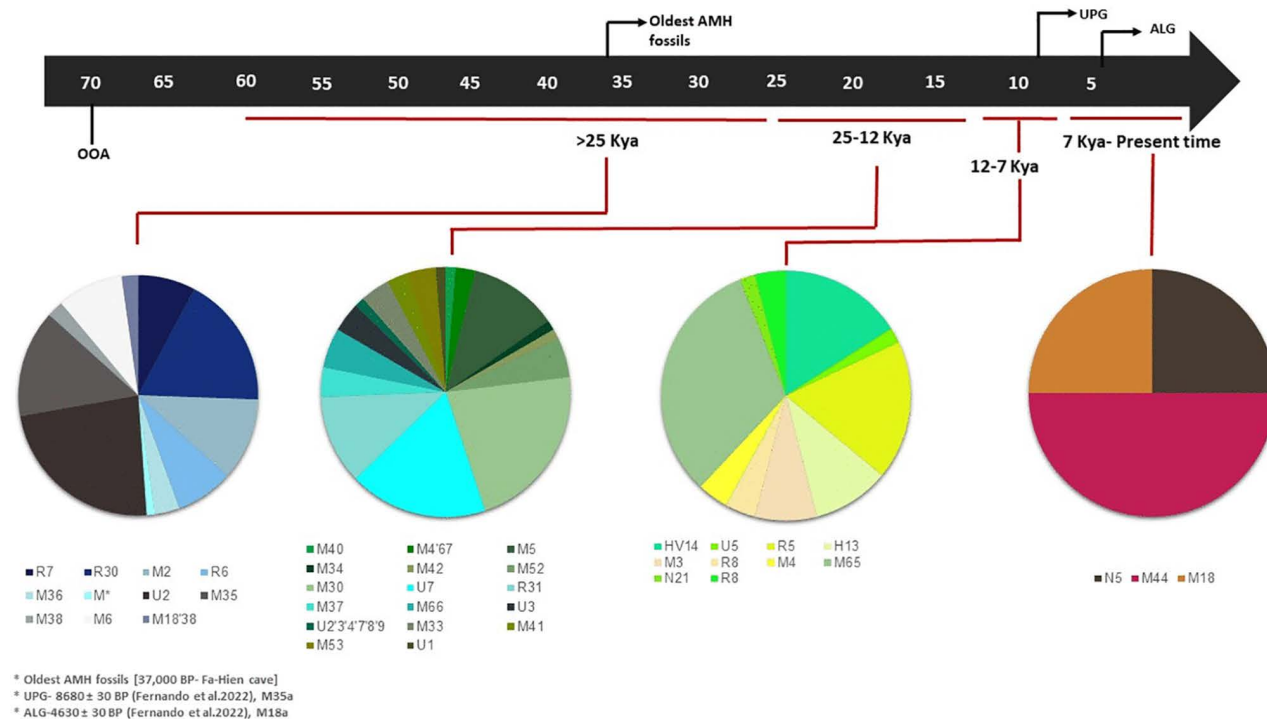
Interestingly, two of these 14 haplogroups (M18 and M35) were observed in ancient individuals analyzed previously from Sri Lanka [40], suggesting long-term maternal continuity. These findings suggest that the maternal genetic heritage of Sri Lanka is largely shaped by South Asian Indigenous populations. These two haplogroups have previously been reported among various tribal groups, including those from West Bengal, Tamil Nadu, Andhra Pradesh, Jharkhand, the Himalayan region, as well as among the Indigenous Tharu population in Nepal [62,63]. We further identified a deep-rooted branch of the rare South Asian haplogroup R7 present in Sri Lanka (Fig 3, S5 Fig, and S4 Table). Although this haplogroup is uncommon, it has a notable frequency among Austroasiatic speakers in India, likely due to later admixture and the founder effect within this group [64]. Surprisingly, R7 was found at high frequencies across Sinhalese and SLT populations [22]. Detailed phylogeographic analysis of the South Asia-specific haplogroup R7 suggested that all the branches identified so far in South Asia are nested within the deeply rooted branches mainly present in the Southern rim of the Indian Ocean (S5 Fig).

The wide yet patchy geographical distribution of haplogroup R7 supports its plausible role as a genetic trace of the Out of Africa migration. The current topology further favors the southern part of South Asia as its likely point of origin, from



**Fig 2. Bayesian Skyline Plots for the Sinhalese, SLTs and Vedda population.**

<https://doi.org/10.1371/journal.pone.0350045.g002>



presence of the East Eurasian haplogroup A1 in the Sri Lankan populations may indicate limited gene flow from East Eurasia to Sri Lanka. In addition, our haplogroup frequency based PCA revealed that all three Sri Lankan populations cluster closely together (S2 Fig) compared to other populations. This pattern is likely attributable to the high degree of haplotype sharing among the study populations (Fig 1). Furthermore, the Sri Lankan groups also exhibit closer genetic affinity to Dravidian populations from parts of South India.

We compared the effective population sizes ( $N_e$ ) of the three major studied populations over time (Fig 2). The SLT population exhibited the largest  $N_e$ , while the Vedda population had the lowest. This indicates that the Vedda population underwent a drastic bottleneck in their maternal genetic histories and, consequently, likely experienced strong genetic drift. In contrast, the SLTs had a higher  $N_e$  compared to the Sinhalese population. Similar observations were reported in a recent whole-genome sequencing study of the Sinhalese, SLTs, and two Vedda clans of Sri Lanka [30]. All three populations show a decline in  $N_e$  during the Holocene (from 12 KYA). The earliest decline was observed in the Vedda population, which may correlate with the time when the land bridge connection between India and Sri Lanka was severed [66]. The sharp decline in  $N_e$  for both the SLTs and Sinhalese is likely related to their later arrival on the island after 5 KYA and the associated founder effect. Moreover, the population expansion of the Vedda occurred slightly later than that of the Sinhalese and SLTs, approximately around 35 KYA. This delayed expansion may be attributed to the concurrent development of microlithic technology on the island, which likely played a role in facilitating the demographic growth of the ancestral Vedda population.

## Conclusions

In conclusion, exploring early human migration and settlement patterns in Sri Lanka reveals a complex population structure, likely reflecting demographic events across different timeframes. Our phylogeographic analysis indicates that the southern route was favored during the initial entry of modern humans into the region, aligning with the genetic data of Asia. The rich heritage of cultural and genetic diversity in Sri Lanka can be traced back to multiple waves of settlement, highlighting the dynamic interactions among various communities over millennia. The analysis of effective population sizes ( $N_e$ ) reveals substantial differences among the three major populations in Sri Lanka, lining up with the lifestyle and subsistence of the populations as well as environmental and demographic factors during the Holocene. The early reduction in the  $N_e$  of Vedda aligns with the severance of connections to India, while declines in both SLTs and Sinhalese may be associated with their arrival on the island in later periods. This research illuminates the role of past demographic processes in shaping the maternal genetic diversity in contemporary Sri Lankan society. Understanding these migrations enhances our knowledge of the island's unique bio-cultural mosaic, as well as drawing connections between ancient and present-day populations.

## Supporting information

**S1 File. Geographic distribution of the maternal haplogroups reported among the Sinhalese, SLTs and Vedda populations.**

(DOCX)

**S1 Fig. Phylogenetic tree construction for mitochondrial haplogroups based on the Maximum Parsimony approach.**

The constructed phylogenetic trees were integrated with data from 144 published sequences belonging to Indian populations and a few other South Asian and global populations. Sinhalese samples are marked in green at the branch ends, while SLTs and Vedda populations are marked in purple and blue, respectively. Samples from the reference populations are displayed in black and are labelled with the name of the population they represent. Each node represents the time to TMRCA with a 95% confidence interval (displayed in red color) which was calculated using Rho statistics.

(XLSX)

**S2 Fig. Principal component analysis (PCA) based on mitochondrial haplogroup frequencies.** This demonstrates that Sri Lankan maternal populations exhibit broad genetic similarity to Indian populations.

(TIF)

**S3 Fig. Mitochondrial haplogroup diversification with time in all three populations.** The temporal classification is based on age estimates obtained from both Rho ( $\rho$ ) estimations and Bayesian skyline analysis.

(TIF)

**S4 Fig. Bayesian phylogenetic tree constructed based on complete mitochondrial genomes of Sinhalese, SLT and Vedda populations.** Based on the coalescent times inferred from the phylogenetic tree, the samples analyzed can be categorized into four distinct temporal periods. The first period corresponds to >25 KYA, the second spans 25–12 KYA, the third encompasses 12–7 KYA and the fourth represents <7 KYA. These periods are visually depicted in the phylogenetic tree using distinct color codes: green, pink, yellow and blue respectively. Additionally, the tip labels were color-coded to differentiate populations, with Sinhalese shown in blue, SLTs in red and the Vedda population in green.

(TIF)

**S5 Fig. Maximum Parsimony tree of R7 haplogroup.** The divergent times were estimated using Rho ( $\rho$ ) statistics.

(TIF)

**S1 Table. Mitochondrial DNA Haplogroup Classification Based on Complete mtDNA Polymorphisms Among Sinhalese, SLTs and Vedda Populations (Sample nos. 1–31: Sanger Sequencing, Current Study; 32–61: NGS, Jayasekara et al, 2023; 62–91: NGS, Current Study; 92–106: Sanger Sequencing, Current Study; 107–123: NGS, Current Study; 124–128: NGS, Fernando et al, 2023; 129–231: NGS, 1000 Genomes Project; 232–237: Sanger Sequencing, Current Study).**

(PDF)

**S2 Table. Accession numbers, haplogroup assignments, and inferred regional origins of published mitochondrial sequences used for phylogenetic tree construction (N = 144).**

(PDF)

**S3 Table. Haplogroup frequencies of different ethnic groups used to plot the PCA.**

(CSV)

**S4 Table. The expansion time of mitochondrial haplogroups based on Rho ( $\rho$ ) age estimates in three Sri Lankan study populations (N = 242).**

(PDF)

## Acknowledgments

We are thankful to the Vedda leaders for their invaluable support and consent, enabling us to conduct sample collection. We are immensely grateful to all the study participants for their voluntary contributions to this study. Furthermore, we thank Venerable Wellampitiye Sumanadhamma and Aludeniye Subodhi Theros, Dr. Nadun Wickramaratne, Mr. K. M. Chamilanga De Silva, Dr. Chamarie Wijesinghe (Medical Officer Incharge, Divisional Hospital, Iccalampattu, Sri Lanka), Dr. Sarath Bandara, MOH staff at Echchalampaththu and Dr. Sonali Randika Premaratne (Post Graduate Institute of Archaeology, University of Kelaniya, Sri Lanka) for their assistance in collecting Ratugala and Echchalampaththu Vedda Samples. The analyses were performed in the High-Performance Computing Centre, Param Shivay, NSCM of IIT, BHU.

\*\* Although the indigenous population of Sri Lanka is currently referred to as 'Adivasi (Vedda),' they were historically

identified as ‘Veddhas’ or ‘Veddhas.’ Since multiple studies have recognized them as a distinct ethnic group, the present study also adheres to the historic term ‘Vedda’.

## Author contributions

**Conceptualization:** Gyaneshwer Chaubey, Ruwandi Ranasinghe.

**Data curation:** Anjana Welikala, Amali Fernando, Joanne Kotelawala, Lakshika Jayasekara, Jose Antonio Urban Aragon, David Witonsky.

**Formal analysis:** Anjana Welikala, Shailesh Desai.

**Investigation:** Anjana Welikala, Ruwandi Ranasinghe.

**Methodology:** Anjana Welikala, Amali Fernando, Joanne Kotelawala, Lakshika Jayasekara, Anoma Jayasoma.

**Project administration:** Ruwandi Ranasinghe.

**Resources:** Ruwandi Ranasinghe.

**Software:** Anjana Welikala, Shailesh Desai, Prajval Pratap Singh, Jose Antonio Urban Aragon, David Witonsky.

**Supervision:** Ruwandi Ranasinghe.

**Visualization:** Anjana Welikala, Shailesh Desai, Prajval Pratap Singh.

**Writing – original draft:** Anjana Welikala.

**Writing – review & editing:** Shailesh Desai, Prajval Pratap Singh, Niraj Rai, Kumaraswamy Thangaraj, Maanasa Raghavan, Gamini Adikari, Kamani H. Tennekoon, Gyaneshwer Chaubey, Ruwandi Ranasinghe.

## References

1. Kulathilake S. The peopling of Sri Lanka from prehistoric to historic times: Biological and archaeological evidence. In: Schug GR, Walimbe SR, editors. *A Companion to South Asia in the Past*. Hoboken (NJ): John Wiley & Sons, Inc. 2016. p. 426–36.
2. Department of Meteorology. Monthly weather summary. [cited 2025 Jan 16] [https://meteo.gov.lk/index.php?option=com\\_content&view=article&id=195&Itemid=565&lang=en](https://meteo.gov.lk/index.php?option=com_content&view=article&id=195&Itemid=565&lang=en)
3. Department of Census and Statistics. Census of population and housing: population atlas of Sri Lanka 2024. Ministry of Finance and Planning, Government of Sri Lanka; 2024.
4. Stringer C. The origin and evolution of Homo sapiens. *Philos Trans R Soc Lond B Biol Sci*. 2016;371(1698):20150237. <https://doi.org/10.1098/rstb.2015.0237> PMID: 27298468
5. James HVA, Petraglia MD. Modern human origins and the evolution of behavior in the later Pleistocene record of South Asia. *Curr Anthropol*. 2005;46(S5):S3–27. <https://doi.org/10.1086/444365>
6. Mellars P. Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. *Proc Natl Acad Sci U S A*. 2006;103:9381–6. <https://doi.org/10.1073/pnas.0510792103>
7. Ghirotto S, Penso-Dolfin L, Barbujani G. Genomic evidence for an African expansion of anatomically modern humans by a Southern route. *Hum Biol*. 2011;83(4):477–89. <https://doi.org/10.3378/027.083.0403> PMID: 21846205
8. Kennedy KAR. To what extent were prehistoric Sri Lankans isolated from the Indian mainland? Biotic and archaeological considerations. In: Braham G, Raczek TP, Rizvi UZ, editors. *Connections and complexity: new approaches to the archaeology of South Asia*. Walnut Creek (CA): Left Coast Press; 2013. p. 35–45.
9. Deraniyagala SU. The prehistory of Sri Lanka: an ecological perspective. Colombo: Department of Archaeological Survey, Government of Sri Lanka; 1992.
10. Kennedy KAR. God-apes and fossil men: Paleoanthropology of South Asia. Ann Arbor: University of Michigan Press; 2000.
11. Wedage O, Picin A, Blinkhorn J, Douka K, Deraniyagala S, Kourampas N, et al. Microliths in the South Asian rainforest ~45–4 ka: new insights from Fa-Hien Lena Cave, Sri Lanka. *PLoS One*. 2019;14(10):e0222606. <https://doi.org/10.1371/journal.pone.0222606> PMID: 31577796
12. Kennedy KAR, Deraniyagala SU. Fossil remains of 28,000-year-old hominids from Sri Lanka. *Curr Anthropol*. 1989;30(3):394–9.
13. Perera N, Kourampas N, Simpson IA, Deraniyagala SU, Bulbeck D, Kamminga J, et al. People of the ancient rainforest: late Pleistocene foragers at the Batadomba-lena rockshelter, Sri Lanka. *J Hum Evol*. 2011;61(3):254–69. <https://doi.org/10.1016/j.jhevol.2011.04.001> PMID: 21777951
14. Kulathilaka S. An evidence-based anthropological exploration of Lanka's people. *Ancient Lanka*. 2022.

15. Allchin B. The late stone age of Ceylon. *J R Anthropol Inst G B Irel*. 1958;88:179–201. <https://doi.org/10.2307/2844251>
16. Deraniyagala PEP. An open-air habitation site of *Homo sapiens balangodensis*. *Spolia Zeylanica*. 1958;28(2):223–61.
17. Kennedy KAR. Human skeletal material from Ceylon with an analysis of the island's prehistoric and contemporary populations. *Bulletin of the British Museum (Natural History) Geology*. 1965;11(4):135–214. <https://doi.org/10.5962/p.313883>
18. Hawkey D. The peopling of South Asia: evidence for affinities and microevolution of prehistoric populations of India and Sri Lanka. In: Wickramasinghe N, editor. *The Department of National Museums*; 2002.
19. Geiger W. *The Mahavamsa or The Greatest Chronicle of Ceylon*. London: Oxford University Press; 1912.
20. Wickremasinghe DMD. *Epigraphia Zelanica; Archaeological Survey of Ceylon*. London: Oxford University Press; 1912.
21. Paranavitana S. *Inscription of Ceylon: early Brahmi inscriptions (Volume I)*. Colombo: The Department of Archaeology Ceylon; 1970.
22. Welikala A, Desai S, Pratap Singh P, Fernando A, Thangaraj K, van Driem G, et al. The genetic identity of the Vedda: a language isolate of South Asia. *Mitochondrion*. 2024;76:101884. <https://doi.org/10.1016/j.mito.2024.101884> PMID: 38626841
23. Illeperuma RJ, Mohotti SN, De Silva TM, Fernandopulle ND, Ratnasooriya WD. Genetic profile of 11 autosomal STR loci among the four major ethnic groups in Sri Lanka. *Forensic Sci Int Genet*. 2009;3(3):e105-6. <https://doi.org/10.1016/j.fsigen.2008.10.002> PMID: 19414153
24. Illeperuma RJ, Markalanda D, Mountain JL, Ratnasooriya WD, Fernandopulle ND, Bamshad MJ. Haplotype data for 12 Y-chromosome STR loci of Sri Lankans. *Forensic Sci Int Genet*. 2010;4(4):e119-20. <https://doi.org/10.1016/j.fsigen.2009.10.003> PMID: 20457052
25. Ranaweera L, Kaewsutthi S, Win Tun A, Boonyarit H, Poolsuwan S, Lertrit P. Mitochondrial DNA history of Sri Lankan ethnic people: their relations within the island and with the Indian subcontinental populations. *J Hum Genet*. 2014;59(1):28–36. <https://doi.org/10.1038/jhg.2013.112> PMID: 24196378
26. Ranasinghe R, Tennekoon KH, Karunanayake EH, Lembring M, Allen M. 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. *Leg Med (Tokyo)*. 2015;17(6):539–46. <https://doi.org/10.1016/j.legalmed.2015.05.007> PMID: 26065620
27. Welikala AHJ, Ranasinghe R, Tennekoon KH, Kotelawala JT, Weerasooriya PR. Mitochondrial DNA (CA)<sub>n</sub> dinucleotide repeat variations in Sinhalese and Vedda populations in Sri Lanka. *Genetica*. 2022;150(2):145–50. <https://doi.org/10.1007/s10709-022-00150-0> PMID: 35141800
28. Perera N, Galhena G, Ranawaka G. X-chromosomal STR based genetic polymorphisms and demographic history of Sri Lankan ethnicities and their relationship with global populations. *Sci Rep*. 2021;11(1):12748. <https://doi.org/10.1038/s41598-021-92314-9> PMID: 34140598
29. Singh PP, Kumar S, Pasupuleti N, Weerasooriya PR, van Driem G, Tennekoon KH, et al. Reconstructing the population history of the Sinhalese, the major ethnic group in Sri Lanka. *iScience*. 2023;26(10):107797. <https://doi.org/10.1016/j.isci.2023.107797> PMID: 37744037
30. Aragon JA, Bandyopadhyay E, Fernando AS, Castro CF, Welikala AHJ, Biddanda A. Whole-genome analysis of Sinhalese and Adivasi sheds light on the human population history of Sri Lanka. *Current Biology*. 2025;35(11):2554–2566.e7. <https://doi.org/10.1016/j.cub.2025.04.039>
31. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2009;30(2):E386–94. <https://doi.org/10.1002/humu.20921> PMID: 18853457
32. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526:68–74. <https://doi.org/10.1038/nature15393>
33. Silva M, Oliveira M, Vieira D, Brandão A, Rito T, Pereira JB, et al. A genetic chronology for the Indian Subcontinent points to heavily sex-biased dispersals. *BMC Evol Biol*. 2017;17(1):88. <https://doi.org/10.1186/s12862-017-0936-9> PMID: 28335724
34. Tiphania Kotelawala J, Ranasinghe R, Rodrigo C, Tennekoon KH, Silva KD. Evaluation of non-coding region sequence variants and mitochondrial haplogroups as potential biomarkers of sporadic breast cancer in individuals of Sri Lankan Sinhalese ethnicity. *Biomed Rep*. 2020;12(6):339–47. <https://doi.org/10.3892/br.2020.1292> PMID: 32346478
35. Rieder MJ, Taylor SL, Tobe VO, Nickerson DA. Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. *Nucleic Acids Res*. 1998;26(4):967–73. <https://doi.org/10.1093/nar/26.4.967> PMID: 9461455
36. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, et al. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290(5806):457–65. <https://doi.org/10.1038/290457a0> PMID: 7219534
37. Okonechnikov K, Golosova O, Fursov M, UGENE team. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*. 2012;28(8):1166–7. <https://doi.org/10.1093/bioinformatics/bts091> PMID: 22368248
38. Kogelnik AM, Lott MT, Brown MD, Navathe SB, Wallace DC. MITOMAP: a human mitochondrial genome database. *Nucleic Acids Res*. 1996;24(1):177–9. <https://doi.org/10.1093/nar/24.1.177> PMID: 8594574
39. Jayasekera LP, Ranasinghe R, Senathilake KS, Kotelawala JT, de Silva K, Abeygunasekara PH, et al. Mitochondrial genome in sporadic breast cancer: a case control study and a proteomic analysis in a Sinhalese cohort from Sri Lanka. *PLoS One*. 2023;18(2):e0281620. <https://doi.org/10.1371/journal.pone.0281620> PMID: 36758048
40. Fernando AS, Wanninayaka A, Dewage D, Karunanayake EH, Rai N, Somadeva R, et al. The mitochondrial genomes of two Pre-historic Hunter Gatherers in Sri Lanka. *J Hum Genet*. 2023;68(2):103–5. <https://doi.org/10.1038/s10038-022-01099-w> PMID: 36450887
41. Kloss-Brandstätter A, Pacher D, Schönherr S, Weissensteiner H, Binna R, Specht G, et al. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. *Hum Mutat*. 2011;32(1):25–32. <https://doi.org/10.1002/humu.21382> PMID: 20960467

42. Sahakyan H, Hooshir Kashani B, Tamang R, Kushniarevich A, Francis A, Costa MD, et al. Origin and spread of human mitochondrial DNA haplogroup U7. *Sci Rep*. 2017;7:46044. <https://doi.org/10.1038/srep46044> PMID: [28387361](#)
43. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*. 2017;34(12):3299–302. <https://doi.org/10.1093/molbev/msx248> PMID: [29029172](#)
44. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010;10(3):564–7. <https://doi.org/10.1111/j.1755-0998.2010.02847.x> PMID: [21565059](#)
45. Eltsov N, Volodko N. mtPhyl: Software tool for human mtDNA analysis and phylogeny reconstruction. 2011.
46. Forster P, Harding R, Torroni A, Bandelt HJ. Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet*. 1996;59(4):935–45. PMID: [8808611](#)
47. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, et al. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 2019;15(4):e1006650. <https://doi.org/10.1371/journal.pcbi.1006650> PMID: [30958812](#)
48. Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, et al. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet*. 2009;84(6):740–59. <https://doi.org/10.1016/j.ajhg.2009.05.001> PMID: [19500773](#)
49. Bandelt H-J, Kloss-Brandstätter A, Richards MB, Yao Y-G, Logan I. The case for the continuing use of the revised Cambridge Reference Sequence (rCRS) and the standardization of notation in human mitochondrial DNA studies. *J Hum Genet*. 2014;59(2):66–77. <https://doi.org/10.1038/jhg.2013.120> PMID: [24304692](#)
50. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 2012;9(8):772. <https://doi.org/10.1038/nmeth.2109> PMID: [22847109](#)
51. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49(W1):W293–6. <https://doi.org/10.1093/nar/gkab301> PMID: [33885785](#)
52. Posit team. RStudio: Integrated Development Environment for R. Boston, MA: Posit Software, PBC; 2025.
53. Anthropological Survey of India. Genomic diversity in people of India, focus on mtDNA and Y-chromosome polymorphism. Singapore: Springer Nature Pvt. Ltd.; 2021.
54. Pagani L, Lawson DJ, Jagoda E, Mörseburg A, Eriksson A, Mitt M, et al. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature*. 2016;538(7624):238–42. <https://doi.org/10.1038/nature19792> PMID: [27654910](#)
55. Desai S, Alqaisi MHM, Singh PP, Pandey P, Pandey RK, Mishra RK, et al. Maternal genetic ancestry of Gujarat: a Pleistocene passage to India. *Am J Biol Anthropol*. 2026;189(1):e70193. <https://doi.org/10.1002/ajpa.70193> PMID: [41474133](#)
56. Sharma I, Sharma V, Khan A, Kumar P, Rai E, Bamezai RNK, et al. Ancient human migrations to and through Jammu Kashmir- India were not of males exclusively. *Sci Rep*. 2018;8(1):851. <https://doi.org/10.1038/s41598-017-18893-8> PMID: [29339819](#)
57. Thangaraj K, Chaubey G, Kivisild T, Reddy AG, Singh VK, Rasalkar AA, et al. Reconstructing the origin of Andaman Islanders. *Science*. 2005;308(5724):996. <https://doi.org/10.1126/science.1109987> PMID: [15890876](#)
58. Mellars P, Gori KC, Carr M, Soares PA, Richards MB. Genetic and archaeological perspectives on the initial modern human colonization of southern Asia. *Proc Natl Acad Sci U S A*. 2013;110(26):10699–704. <https://doi.org/10.1073/pnas.1306043110> PMID: [23754394](#)
59. Gakuhari T, Nakagome S, Rasmussen S, Allentoft ME, Sato T, Korneliusson T, et al. Ancient Jomon genome sequence analysis sheds light on migration patterns of early East Asian populations. *Commun Biol*. 2020;3(1):437. <https://doi.org/10.1038/s42003-020-01162-2> PMID: [32843717](#)
60. Jha G, Sinha DK, Jha DK, Ajithprasad P. Coasting into India? - Assessing lithostratigraphic context of Middle Palaeolithic occupation in Saurashtra Peninsula. *Quat Environ Hum*. 2024;2(6):100034. <https://doi.org/10.1016/j.qeh.2024.100034>
61. Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, et al. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*. 2005;308(5724):1034–6. <https://doi.org/10.1126/science.1109792> PMID: [15890885](#)
62. Chandrasekar A, Kumar S, Sreenath J, Sarkar BN, Urade BP, Mallick S, et al. Updating phylogeny of mitochondrial DNA macrohaplogroup m in India: dispersal of modern human in South Asian corridor. *PLoS One*. 2009;4(10):e7447. <https://doi.org/10.1371/journal.pone.0007447> PMID: [19823670](#)
63. Fornarino S, Pala M, Battaglia V, Maranta R, Achilli A, Modiano G, et al. Mitochondrial and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation. *BMC Evol Biol*. 2009;9:154. <https://doi.org/10.1186/1471-2148-9-154> PMID: [19573232](#)
64. Chaubey G, Karmin M, Metspalu E, Metspalu M, Selvi-Rani D, Singh VK, et al. Phylogeography of mtDNA haplogroup R7 in the Indian peninsula. *BMC Evol Biol*. 2008;8:227. <https://doi.org/10.1186/1471-2148-8-227> PMID: [18680585](#)
65. Larruga JM, Marrero P, Abu-Amro KK, Golubenkov MV, Cabrera VM. Carriers of mitochondrial DNA macrohaplogroup R colonized Eurasia and Australasia from a Southeast Asia core area. *BMC Evol Biol*. 2017;17(1):115. <https://doi.org/10.1186/s12862-017-0964-5> PMID: [28535779](#)
66. Dubey KM, Chaubey AK, Gaur AS, Joglekar MV. Evolution of Ramasetu region as a link between India and Sri Lanka during the late Pleistocene and Holocene. *Quat Res*. 2023;111:166–76. <https://doi.org/10.1017/qua.2022.41>