

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/320507562>

# Paternal lineage affinity of the Malay subethnic and Orang Asli populations in Peninsular Malaysia

Article in *International Journal of Legal Medicine* · July 2018

DOI: 10.1007/s00414-017-1697-0

CITATIONS

4

READS

387

6 authors, including:



**Sharifahanny RahayuKarmilla Syedhassan**

Universiti Sains Malaysia

5 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)



**Aedrianee REEZA Alwi**

Universiti Teknologi Malaysia

6 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)



**Mohd Nor Norazmi**

Universiti Sains Malaysia

179 PUBLICATIONS 1,946 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Recombinant BCG [View project](#)



Research-Development of vaccines and therapies against TB [View project](#)

# Paternal lineage affinity of the Malay subethnic and Orang Asli populations in Peninsular Malaysia

SyedHassan SharifahNany RahayuKarmilla<sup>1</sup> · Alwi R. Aedrianee<sup>2</sup> · Abd Rashid Nur Haslindawaty<sup>1</sup> · Abdullah Nur Azeelah<sup>1</sup> · Sundararajulu Panneerchelvam<sup>1</sup> · Mohd Nor Norazmi<sup>1</sup> · Zainuddin Zafarina<sup>1,3</sup>

Received: 26 April 2017 / Accepted: 20 September 2017  
© Springer-Verlag GmbH Germany 2017

**Abstract** Peninsular Malaysia is populated by the Malays, Chinese, Indians, and Orang Asli. We have analyzed 17 Y-STRs loci for 243 randomly unrelated individuals, which include 153 Malays (7 Acheh, 13 Champa, 11 Rawa, 9 Kedah, 23 Minang, 15 Bugis, 43 Kelantan, 14 Jawa, and 18 Bugis) and 90 Orang Asli [54 Semang (16 Kensiu, 13 Lanoh, 25 Bateq); 30 Senoi (21 Semai, 9 Che Wong); and 6 Proto-Malay (6 Orang Kanaq)] from selected settlements in Peninsular Malaysia using the AmpFISTR Yfiler™ kit (Applied Biosystems™). The overall haplotype diversity is 0.9966, i.e., 0.9984 for the Malays and 0.9793 for the Orang Asli. A total of 158 haplotypes (65.02%) were individually unique. The *p* value and pairwise *R*<sub>st</sub> analysis was calculated to show the genetic structure of the samples with other world populations (from YHRD website). Based on the Y-STR data, Champa, Acheh, Kedah, Minang, and Kelantan are clustered together. Lanoh and Kensiu (Semang) are very closely related, suggesting similar paternal ancestry. Jawa Malays and Indonesian Java, plus the Bugis Malays and Australian Aborigines shared high degree of paternal lineage affinity.

This study presents data for very precious relict groups, who are the earliest inhabitants of Peninsular Malaysia.

**Keywords** Y-STR · Malay · Orang Asli · Peninsular Malaysia

Peninsular Malaysia is populated by the Malays (majority 54.6%), Chinese (24.6%), Indians (7.3%), and Orang Asli (~ 0.5%) (Population and Housing Census of Malaysia, 2010). Malay is termed as an ethnic group of Austronesian peoples predominantly inhabiting Peninsular Malaysia, east Sumatra, southern Thailand, the Tanintharyi region of Myanmar, coastal Borneo including Brunei and West Kalimantan, coastal Sabah and Sarawak [1]. The Malay subethnic groups include Acheh, Champa, Rawa, Kedah, Minang, Bugis, Kelantan, Jawa, Banjar, and Pattani [2]. According to the population survey by Department of Orang Asli Development (JAKOA) in 2011, there are a total of 178,197 Orang Asli individuals in Peninsular Malaysia. The Orang Asli of Peninsular Malaysia is categorized into three major groups—Semang (Negrito phenotype), Senoi, and Proto-Malays and is further divided into six subgroups based on their physical characteristics, linguistic affinities, and cultural and economic practices [3–5]. Semang (short-stature, broad-nose, and frizzy hair) consists of Bateq, Jahai, Kensiu, Kintak, Lanoh, and Mendriq subgroups while Senoi (lighter skin color and mixed Negrito-Mongoloid physical appearance) comprises Che Wong, Jahut, Mah Meri, Semai, Semoq Beri, and Temiar. The Proto-Malays (characteristics almost similar to Malays) includes Jakun, Orang Kanaq, Orang Kuala, Orang Seletar, Semelai, and Temuan. The purpose of this study is to provide population data of these autochthonous Orang Asli subgroups for forensically relevant genetic markers used in forensic and paternity testing. Various Malay and Orang Asli settlements throughout Peninsular Malaysia (north, central, and south

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00414-017-1697-0>) contains supplementary material, which is available to authorized users.

✉ Zainuddin Zafarina  
zafarina@usm.my

<sup>1</sup> Human Identification/DNA Unit, School of Health Sciences, Universiti Sains Malaysia, Health Campus, 11650 Kubang Kerian, Kelantan, Malaysia

<sup>2</sup> Department of Chemistry Malaysia, Jalan Abdul Samad, 80100 Johor Bahru, Johor, Malaysia

<sup>3</sup> Analytical Biochemistry Research Centre (ABrC), Universiti Sains Malaysia, 11800 USM Penang, Malaysia

regions) were chosen. We focused on the groups that are very small in number, least studied and on the verge of “extinction,” i.e., Kensiu, Lanoh, Che Wong, and Orang Kanaq. Samples were also collected from the larger groups (Bateq and Semai) for comparison purpose.

Informed consents were obtained from each participant before the blood samples were collected. Ethical approvals were obtained from Universiti Teknologi MARA and Universiti Sains Malaysia Human Ethics Committee (FWA Reg. No: 00007718; IRB Reg. No: 00004494) with Ref. No: USMKK/PPPJEPeM [254.4.129/]. Genomic DNA was extracted for 243 randomly unrelated individuals, which include 153 Malays (7 Acheh, 13 Champa, 11 Rawa, 9 Kedah, 23 Minang, 15 Bugis, 43 Kelantan, 14 Jawa, and 18 Bugis) and 90 Orang Asli [54 Semang (16 Kensiu, 13 Lanoh, 25 Bateq); 30 Senoi (21 Semai, 9 Che Wong); and 6 Proto-Malay (6 Orang Kanaq)] using QIAamp® DNA blood mini kit (QIAGEN) and amplified using the AmpFISTR Yfiler™ kit (Applied Biosystems™) according to the manufacturer’s protocol. The fragments were separated using 3130xl Genetic Analyzer (Applied Biosystems™) and raw data was analyzed using GeneMapper ID v3.2 Software (Applied Biosystems™). The experimental procedures were carried out in an accredited laboratory (MS ISO 17025:2005) for paternity testing.

Haplotype frequencies and diversity were calculated by GeneA1Ex 6.502 [6]. The haplotype frequency was calculated using the following formula  $F = i/n$ , where  $F$  is the frequency of a haplotype,  $i$  is the fraction of the number of copies, and  $n$  is the population size. Haplotype diversity (HD) was estimated by applying Nei’s formula [7] and was calculated for different combinations of Y-STR loci with the formula  $H = (n/(n-1)) (1 - \sum p_i^2)$ , where  $n$  is the population size and  $p_i$  is the frequency of the  $i$ th haplotype [8]. Haplotype discrimination capacity (DC) was calculated by dividing the total number of different haplotypes in analyzed population by the total number of individuals in the sample size [9].

Haplotype data for 17 Y-STR loci from 153 Malays and 90 Orang Asli males were determined and presented in Table 1. The datasets were submitted to the YHRD after passing the QC requirements for population genetic data and the accession numbers for each subgroup were received: Yan/Sungai Petani/Kedah, Malaysia [Malay-Acheh] (YA004032), Kelantan, Malaysia “Malay-Champa” (YA004033), Perak, Malaysia “Malay-Rawa” (YA004034), Kedah, Malaysia “Malay-Kedah” (YA004035), Negeri Sembilan, Malaysia “Malay-Minang” (YA004036), Johor, Malaysia “Malay-Bugis” (YA004037), Kelantan, Malaysia “Malay-Kelantan” (YA004038), Johor, Malaysia “Malay-Jawa” (YA004039), Johor/Perak, Malaysia “Malay-Banjar” (YA004040), Baling, Kedah, Malaysia “Orang Asli-Kensiu” (YA004133), Lenggong, Perak, Malaysia “Orang Asli-Lanoh” (YA004134), Gua Musang, Kelantan, Malaysia “Orang Asli-Bateq” (YA004131), Kuala Lipis, Pahang, Malaysia “Orang

Asli-Semai” (YA004135) and Lanchang, Pahang, Malaysia “Orang Asli-Che Wong” (YA004132). The forensic efficiency values such as haplotype diversity and discrimination capacity for each Malay and Orang Asli groups are listed in Table 2. The Y-STR haplotypes for the Malay subethnic groups are highly variable and polymorphic, which is in accordance with the previously published data for the random Malay population [10].

A comparative analysis was carried out with 19 other population samples, to give a total of 3444 haplotypes by analysis of molecular variance (AMOVA) and visualized in multidimensional scaling (MDS) plot, using an online tool of the YHRD [11]. The samples were from the following areas (the YHRD designation is given in parentheses): Cape Town, South Africa (“Xhosa”,  $n = 100$  haplotypes); Bagandu, Central African Republic (“Biaka Pygmy”,  $n = 33$  haplotypes); Australia (“Aboriginal”,  $n = 747$  haplotypes); Ketou, Benin (“Yoruba”,  $n = 54$  haplotypes); China (“Han”,  $n = 584$  haplotypes); Southern India, India (“Tamil”,  $n = 122$  haplotypes); Java, Indonesia (“Indonesian”,  $n = 136$  haplotypes); Webuye, Kenya (“Luhya”,  $n = 42$  haplotypes); Kinyawa, Kenya (“Maasai”,  $n = 99$  haplotypes); Sarawak, Malaysia (“Bidayuh”,  $n = 113$  haplotypes; “Iban”,  $n = 96$  haplotypes); Malaysia (“Indian”,  $n = 293$  haplotypes); Ibadan, Nigeria (“Yoruba”,  $n = 112$  haplotypes); Taiwan (“Ami”,  $n = 84$  haplotypes; “Atayal”,  $n = 240$  haplotypes; “Bunun”,  $n = 36$  haplotypes); Northern Thailand, Thailand (“Lawa”,  $n = 12$  haplotypes; “Lua”,  $n = 26$  haplotypes) and Vietnam (“Mong”,  $n = 59$  haplotypes). Significance was set at  $p < 0.05$ . The pairwise  $R_{st}$  value and associated  $p$  values, with 10,000 permutations, between the nine Malay subethnic groups and six Orang Asli subgroups with the world populations are shown in Table 3 and the MDS plot is presented in Fig. 1.

Based on the MDS plot clustering, Champa, Acheh, Kedah, Minang, and Kelantan Malays were clustered together, suggesting close genetic relationship. Interestingly, previous study on SNPs genotypic data had shown that Kelantan Malays form a different cluster from other Malay subethnic groups [12]. Nonetheless, the Y-STR data may suggest similar paternal ancestry of Kelantan and other Malay subethnic groups. Jawa Malays are plotted on the same position with the Indonesian Java ( $p = 0.7053$ ,  $R_{st} = -0.0139$ ), suggesting a common ancestry. The link between the Javanese and Jawa Malays was important during migration in colonial period, spreading of Islam in Indonesia, when religious missionaries were sent from Malacca to seaports on the northern coast of Java [13]. Based on the Y-STR plot, the Bugis seems to be more distant from the rest of the Malay subethnic groups but genetically closer to the Australian Aborigines. The Buginese sailors have traveled and left their mark and culture in the northern Australian coast, as well as having trading activities with the Orang Asli in Peninsular Malaysia [14]. These have resulted into gene flow between these populations and alter the mainstream gene pool. [15]. The Malaysian Genetic Layer

Cake has three basic tiers (i.e., the Orang Asli groups; Semang, Senoi, and Proto-Malays) and two toppings (Malay subethnic groups plus worldwide immigrants). The foundation is formed by ancient Semang people followed by the Senoi who form the second tier [16] while the third layer is made up of admixed Austronesian-speaking people [17].

Furthermore, our data also showed significant differences between the Orang Asli subgroups and other world reference populations as most of the  $p$  values were 0.0000 and  $R_{st}$  values  $> 0.05$ . As expected, no significant genetic distance was observed for Lanoh and Kensiu ( $p = 0.4543$  and  $R_{st} = -0.0196$ ). Both Lanoh and Kensiu belong to Semang and speak Aslian language. The Lanoh and Kensiu settlements are located relatively near to each other in the northern region of Peninsular Malaysia. Both groups were plotted on the same point (Fig. 1), suggesting high degree of affinity between their paternal lineages. This might indicate that they were originally the same group of people and only recently being grouped by the government into smaller subgroups. The close genetic affinity between the Semang and Australian Aborigines as in the MDS plot supports the historical viewpoint on the origin and migration pattern of the male population groups. The Semang (Lanoh, Kensiu, and Bateq), also known as Negrito, is the smallest Orang Asli group and was thought to have originated from Africa spreading through Southeast Asia before colonizing the Australasian region, with the time depth of 60,000 years ago [18]. Comparatively, the Senoi is the largest group of Orang Asli consisting about 54% of the population [19] and totalling around 79,156 individuals [20]. The ancestors of the Senoi were believed to be the second group of peoples to arrive in Peninsular Malaysia during the second wave of migration about 8000 years ago [21]. The Senoi speaks Austroasiatic languages of the Mon-Khmer subgroup, i.e., Semai speaks Central Aslian while Che Wong is the Northern Aslian speakers similar to the Semang. According to the MDS plot, the Semai was displaced toward Southeast Asian population. However, Che Wong was casted out might be due to the sample size. The Senoi were proposed to be originated from Indochina region around 8000 years ago during the Neolithic era and migrated south from mainland of Southeast Asia to Peninsular Malaysia [22]. The Proto-Malays are the descendants of Austronesian-speaking voyagers who migrated from Taiwan through the Philippines to Borneo, Sulawesi, Java and finally settled in the coastal region of Peninsular Malaysia between ~ 6000 and ~ 2500 years ago [23]. Based on the Y-STR plot, Orang Kanaq (representative of Proto-Malays) is much closer to the Taiwanese and Chinese and in fact they were proposed as the first wave of Mongoloid population expansion into the Islands of Southeast Asia [24].

Future work should include haplotyping of more Orang Asli individuals in Peninsular Malaysia as their genetic information is very important to be recorded, not only for population genetic studies but also as a world heritage. The previous

migration events of modern human could be deduced by understanding the genetic makeup of these relict populations.

In conclusion, the 17 Y-STR markers offer a discriminative tool for genetic affinity and population studies for the Malays and Orang Asli. The Y-STR data indicate that the Champa, Aceh, Kedah, Minang, and Kelantan are closely related. The Lanoh and Kensiu may share similar paternal ancestor, in agreement with their physical appearances and language similarities. The Jawa Malays and Indonesian Java along with Bugis Malays and Australian Aborigines shared high degree of paternal lineage affinity. It is very important to report the data of the Orang Asli subgroups as these precious relict groups' population number is not showing significant growth for the past 10 years and may just lost in time. As the earliest settlers of Peninsular Malaysia, it is very important to understand their genetic makeup in order to infer the migration history of Peninsular Malaysia and Southeast Asia.

This paper follows the recommendations of the ISFG on the use of Y-STRs in forensic analysis and the guidelines for publication of population data requested by the journal [25].

**Acknowledgements** The authors acknowledge the Department of Orang Asli Development (JAKOA), Ministry of Higher Education Malaysia and Universiti Sains Malaysia (Long-Term Research Grant Scheme: 304/PPSK/6150115/U132) (Orang Asli samples), and FRGS Top Down 203/PPSK/ 249 6171005 (Malay samples).

**Compliance with ethical standards** Informed consents were obtained from each participant before the blood samples were collected. Ethical approvals were obtained from Universiti Teknologi MARA and Universiti Sains Malaysia Human Ethics Committee (FWA Reg. No: 00007718; IRB Reg. No: 00004494) with Ref. No: USMCK/PPPJEPeM [254.4.129/].

## References

1. Milner A (2010) The Malays (the peoples of South-East Asia and the Pacific). John Wiley and Sons Ltd, United Kingdom
2. Halim-Fikri H, Etemad A, Latif AZA, Merican AF, Baig AA, Annuar AA, Ismail E, Salahshourifar I, Liza-Sharmini AT, Ramli M, Shah MI, Johan MF, Hassan NNN, Abdul-Aziz NM, Noor NHM, Nur-Shafawati AR, Hassan R, Bahar R, Zain RB, Yusoff SM, Yusoff S, Tan SG, Thong M-K, Wan-Isa H, Abdullah WZ, Mohamed Z, Latiff ZA, Zilfalil BA (2015) The first Malay database toward the ethnic-specific target molecular variation. BMC Res Notes 8(1):176. <https://doi.org/10.1186/s13104-015-1123-y>
3. Bellwood P (1997) Prehistory of the Indo-Malaysian Archipelago. Univ. of Hawaii Press, Honolulu
4. Nicholas C (2006) In: Hood S (ed) The Orang Asli: origins, identity and classification. Archipelago Press, Kuala Lumpur, pp 20–21
5. JHEOA (2002) Kehidupan, budaya dan pantang larang orang asli. Department of Orang Asli, Kuala Lumpur
6. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295
7. Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

8. Kayser M, Knijff PD, Dieltjes P, Kraweczek M, Nagy M, Zerjal T et al (1997) Applications of microsatellite-based Y-chromosome haplotyping. *Electrophoresis* 18:1602–1607
9. Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G et al (1997) Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* 110:125–133 41–49
10. Chang YM, Perumal R, Keat PY, Kuehn DLC (2007) Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. *Forensic Sci Int* 167:70–76
11. Willuweit S, Roewer L (2015) The new Y chromosome haplotype reference database. *Forensic Sci Int Genet* 15:43–48
12. Hatin WI, Nur-Shafawati AR, Zahri M-K, Xu S, Jin L, Tan S-G et al (2011) Population genetic structure of Peninsular Malaysia Malay sub-ethnic groups. *PLoS One* 6:e18312
13. Wink A (2004) *Indo-Islamic society: 14th–15th centuries*. Oxford University Press, Oxford, England, UK
14. Schwerdtner Máñez K, Ferse SCA (2010) The history of Makassan trepang fishing and trade. *PLoS ONE* 5(6):e11346
15. Norhalifah HK, Syaza FH, Chambers GK, Edinur HA (2016) The genetic history of Peninsular Malaysia. *Gene* 586:129–135
16. Bellwood P (2005) *First farmers: the origins of agricultural societies*. Blackwell, Victoria
17. Tumonggor MK, Karafet TM, Hallmark B, Lansing JS, Sudoyo H, Hammer MF et al (2013) The Indonesian archipelago: an ancient genetic highway linking Asia and the Pacific. *J Hum Genet* 58: 165–173
18. Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W et al (2005) Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308:1034
19. Nicholas C (2000) *The Orang Asli and the contest for resources: indigenous politics, development and identity in Peninsular Malaysia*. International Work Group on Indigenous Affairs (IWGIA), Denmark
20. JHEOA (2008) *Data Maklumat Asas Jabatan Hal Ehwal Orang Asli Tahun. Bahagian Penyelidikan dan Perancangan Jabatan Hal Ehwal Orang Asli*, Kuala Lumpur
21. Baer A (1999) *Health, disease and survival: a biomedical and genetic analysis of the Orang Asli of Malaysia* COAC, Kuala Lumpur
22. Hill C, Soares P, Mormina M, Macaulay V, Meehan W, Blackburn J, Clarke D, Raja JM, Ismail P, Bulbeck D, Oppenheimer S, Richards M (2006) Phylogeography and ethnogenesis of aboriginal southeast Asians. *Mol Biol Evol* 23(12):2480–2491
23. Andaya LY (2001) The search for the ‘origins’ of Melayu. *J Southeast Asian Stud* 32:315–330
24. Chang YM, Swaran Y, Phoon YK, Sothirasan K, Sim HT, Lim KB et al (2009) Haplotype diversity of 17 Y-chromosomal STRs in three native Sarawak populations (Iban, Bidayuh and Melanau) in East Malaysia. *Forensic Sci Int Genet* 3:e77–e80
25. Poetsch M, Bajanowski T, Pfeiffer H (2012) The publication of population genetic data in the International Journal of Legal Medicine: guidelines. *Int J Legal Med* 126:489–490