

Research Note

CHARACTERIZATION OF SEVEN (CA)_n MICROSATELLITE MARKERS AMONG THREE POPULATIONS OF NEGRITOS FROM PENINSULAR MALAYSIA

HOH, B.P.^{1,2,*}, NUR SHAFAWATI, A.R.¹ and YAM, Y.Y.¹

¹Human Genome Centre, School of Medical Sciences, Health Campus,
Universiti Sains Malaysia, Kota Bharu, Kelantan, Malaysia

²Institute of Medical Molecular Medicine, Faculty of Medicine, Universiti Teknologi MARA,
Sungai Buloh Campus, Selangor, Malaysia

*E-mail: hohpeng@yahoo.com

Southeast Asia has the longest record of human habitation outside of Africa, and occupied a pivotal role for nearly two million years following the arrival of *Homo erectus*, in what is now Java (Jin *et al.*, 2001). The Negritos, believed to be the earliest aborigines arrived in Peninsular Malaysia dating more than 10,000 years ago (Bellwood, 1993), are thought to be related to Andaman Islanders and the aborigines of the Philippines. In the peninsula, they are divided into six groups namely, Bateq, Mendriq, Jahai, Kensiu, Lanoh and Kintak, based in their own linguistic, socio-cultural and geographical regions.

Microsatellite markers have key roles in various fields of genetics such as forensic genetics, anthropological genetics, genealogy and evolutionary genetics because they are highly variable in repeat lengths and polymorphisms hence more informative. In addition, microsatellites typically having more than 10 alleles as compared to other types of markers (usually have two or three per locus), offering a great statistical power for data analysis. Despite the advancement of high throughput SNP genotyping technology, data on microsatellite markers remains widely reported (eg, Fernandes *et al.*, 2003; Park *et al.*, 2008; Tishkoff *et al.*, 2009). Early population genetic studies on the aborigines in particular the Negritos in Peninsular Malaysia were carried out using classical genetic markers (Tan, 1979; Tan & Teng, 1978); and later with maternally inherited mtDNA markers (Hill *et al.*, 2006; Macaulay *et al.*, 2005). However, study of autosomal microsatellite markers on this native population has yet to be reported extensively.

Our preliminary findings suggest the possibility of inbreeding and genetic drift in the three Negrito

tribes from Northern Peninsular Malaysia namely, Jahai, Bateq and Mendriq. In this paper, we report this finding in a total of seven randomly selected highly polymorphic dinucleotide (CA)_n microsatellite markers. We observed that the Negrito population showed a high level of genetic similarity among the tribes but a significantly lower observed heterozygosity when compared to other population in the published data (CEPH database; Tan *et al.*, 2002).

The study was approved by the USM ethics committee and official approval was obtained from the Department of Orang Asli Development Malaysia (*Jabatan Kemajuan Orang Asli, JAKOA*). Informed written consent, which was conducted in Malay or local dialect, and accompanied by an officer from the Department of Orang Asli Development, was obtained from all participating individuals.

Stringent inclusion and exclusion criteria were applied during sample collection. Subject must be i) descended from the same tribe from the same location for at least three generations; ii) healthy and free from any known systemic diseases. Subjects were excluded if they i) were descended from different ethnic group or tribe within three generations; ii) unhealthy or carrying any known systemic diseases. Interviews were carried out and pedigree information was obtained during sampling. We observed a high degree of consanguineous and/or inter-ethnic marriages among these populations (data not shown). For this reason, only "unrelated" individuals (defined as non-consanguineous marriage; at least second degree relatives) were selected; while the remaining were excluded from further analyses, resulting in 62 unrelated individuals from three Negrito tribes namely Jahai, Mendriq and

* To whom correspondence should be addressed.

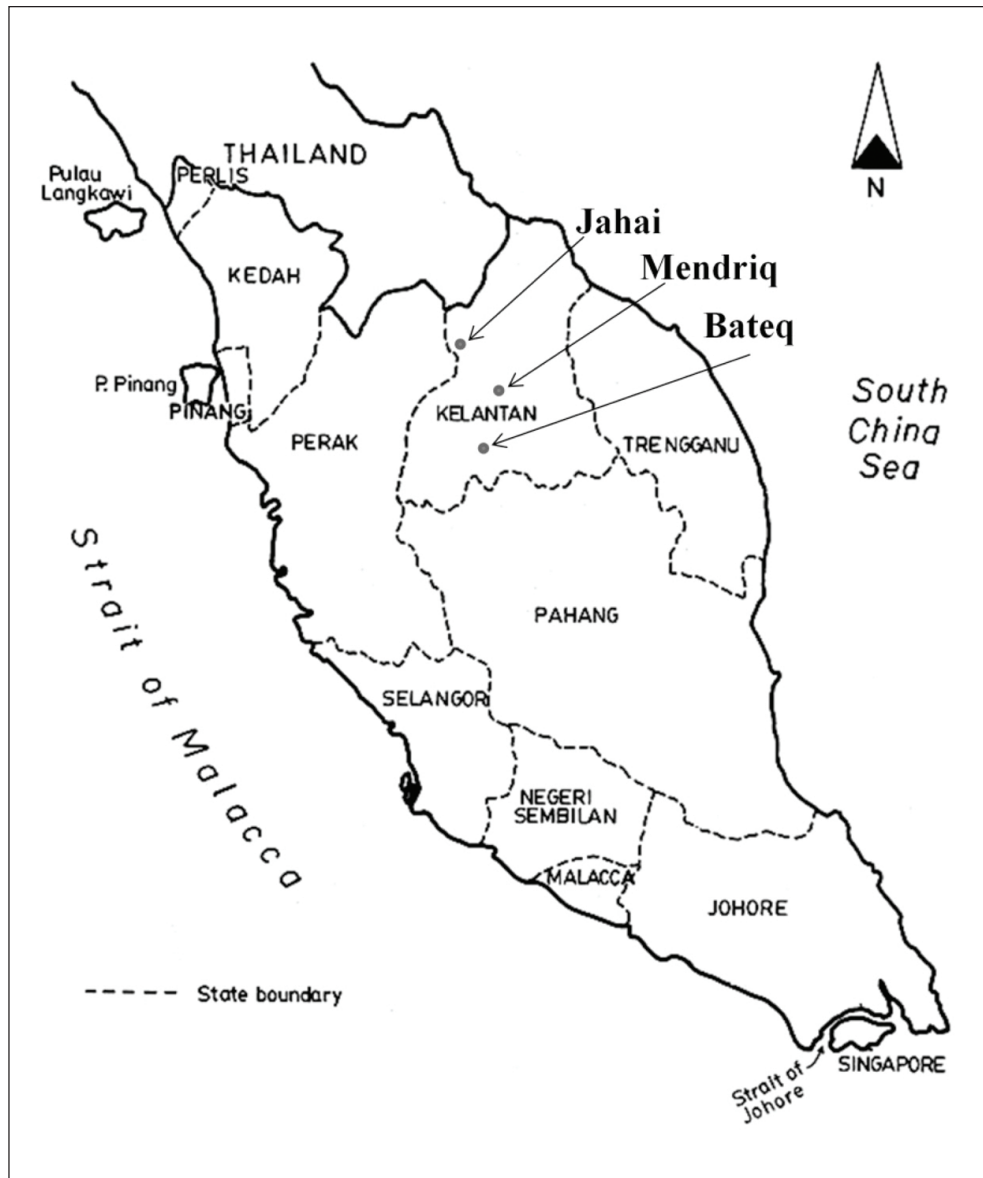


Fig. 1. Map of Peninsular Malaysia indicating the sampling locations of the three Negrito tribes, namely Jahai, Mendriq and Bateq in the state of Kelantan.

Bateq. The sampling locations are shown in Figure 1. Three milliliter of venous blood was taken from the subjects, and DNA was extracted from the whole blood.

Seven CA repeat motif microsatellite primers, with heterozygosities and polymorphic information content (PIC) more than 0.75, were selected randomly from Tan *et al.* (2002) and the CEPH Genotype Database (<http://www.cephb.fr/cephdb/php/>) (Table 1). Optimization and amplification of PCR were carried out in a total volume of 25 ml reaction mixture consisting of 22.5 ml of ABGene Ready Mix (Thermo Fisher Scientific, Surrey, UK), 0.2 μ M of each microsatellite primer pairs and ~10 ng of template DNA. Presence of amplicons were checked by 4% agarose gel (Metaphor, Cambrex,

Bioscience, Rockland Maine, US) electrophoresis and then diluted in the ratio of 1:20. Fluorescent-labeled fragments were electrophoresed through ABI3100 and analyzed using GeneScan Analysis (Ver 2.0) (Figure 2). The allelic frequencies for each marker were computed by counting the alleles in unrelated individuals. The observed and expected heterozygosities as well as Hardy-Weinberg Equilibrium (HWE) tests were computed. Marker information for Caucasian and Chinese were obtained from the CEPH database and published data (Tan *et al.*, 2002).

Subjects with ambiguous genotypes or unsuccessful PCR amplification after three attempts were excluded from the analysis. The number of alleles ranged from eight (D13S153) to 13

Table 1. Characteristics of the seven microsatellite loci examined in Negritos from Northern Peninsular Malaysia and comparisons of observed heterozygosity values with the CEPH and Chinese populations.

Locus	Primer sequence	Expected allele sizes, bp (Repeat length)	Observed no of allele	H _e (H _o)	*H _{Chinese}	*H _{CEPH}
D1S243	F: CACACAGGCTCACATGCC - HEX R: GCTCCAGCGTCATGGACT	162 (AC) ₂₄	11	0.8772 (0.8545)	-	0.8148
D3S1298	F: AGCTCTCAGTGCCACCC - 6FAM R: GAAAAATCCCCTGTGAAGCG	200 (AC) ₂₆	10	0.8554 (0.3261)	0.91	0.8889
D7S517	F: TGGAGAAGCCATGTGAGT - 6FAM R: AGCTGTAATTAGTTGCTGGTTGA	255 (AC) ₂₆	9	0.7969 (0.6607)	-	0.9259
D8S1706	F: AGACAGCGACATGCAG - 6FAM R: GAAAAATCACAATAGAAATTACT	298 (CA) ₂₁	13	0.7396 (0.5400)	-	0.8929
D12S339	F: TCAGGGTTTGAGACTNGCC - HEX R: GTCACGCCCTGGCTAGTGT	269 (AC) ₃₀	10	0.8368 (0.2708)	0.89	0.8519
D13S153	F: AGCATTGTTTCATGTTGGTG - 6FAM R: CAGCAGTGAAGGCTAAGCC	222 (CA) ₂₅	8	0.7779 (0.4182)	0.90	0.8100
D19S220	F: ATGTTCAGAAAGGCCATGTCATTG - HEX R: TCCCTAACGGATACACAGCAACAC	275 (AC) ₂₇	13	0.8345 (0.2885)	0.85	0.9286

Expected heterozygosity values were computed using Nei's (1973)

H_o, Observed heterozygosity; H_e, expected heterozygosity; H_{Chinese}, heterozygosity of Chinese population; H_{CEPH}, heterozygosity of Caucasian population

* Information was obtained from either published data (Tan *et al.*, 2002) or CEPH Genotype Database (CEPH, <http://www.cephb.fr/cephdb/>)

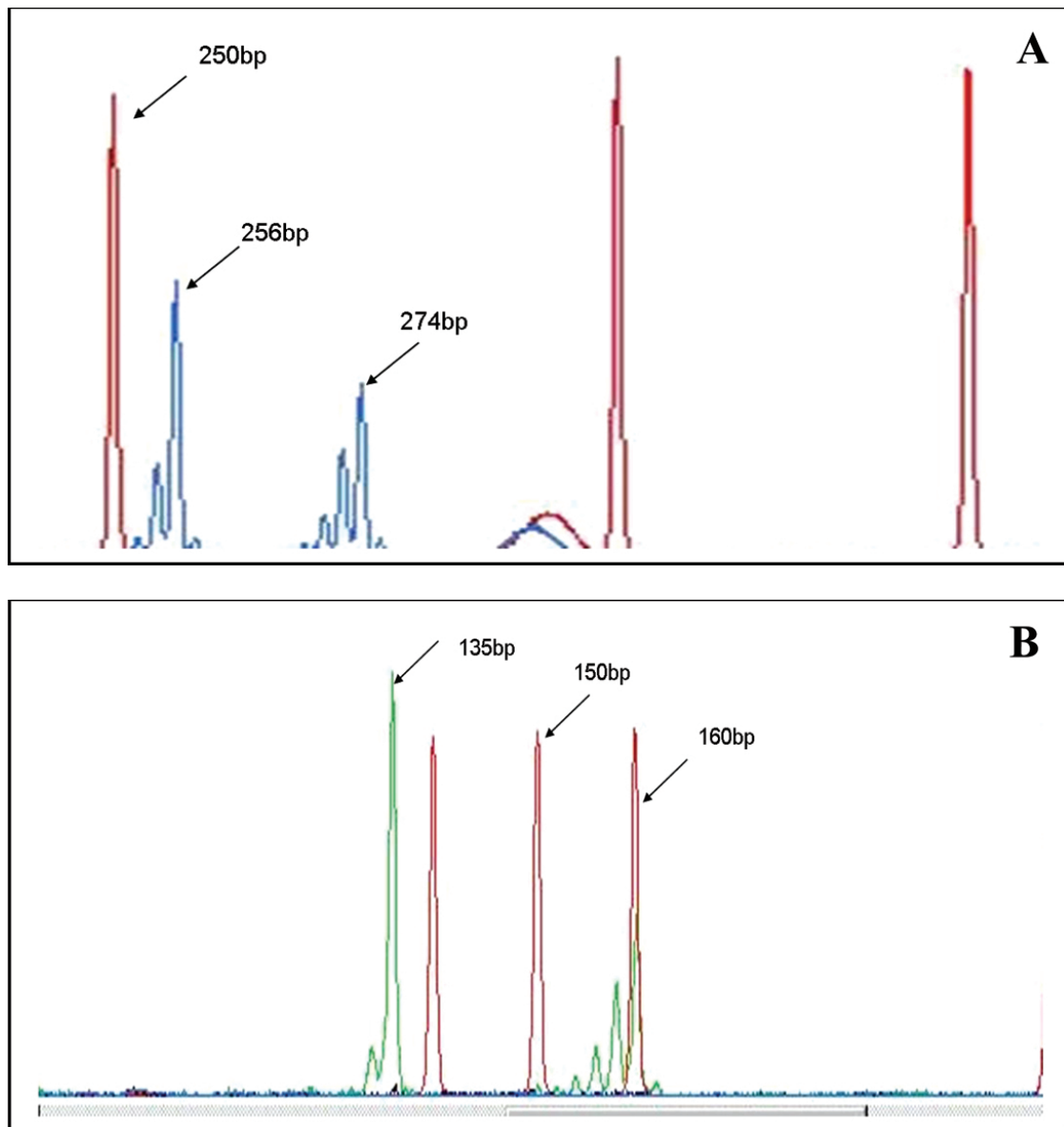


Fig. 2. Examples of electrophoregrams showing heterozygote peaks of a Mendriq individual M14 at locus (A) D8S1706 (256bp / 274 bp); and (B) D1S243 (135 bp / 160 bp). The red peaks indicate the size standard labeled with fluorescence dye ROX: (A) 250 bp; (B) 150 bp and 160 bp.

(D8S1708) (Table 1); while heterozygosities of the seven loci in the Negrito populations were calculated and summarized in Table 1. Comparisons of heterozygosities of the Negritos with the available data for Caucasian and Chinese populations are presented in Figure 3. The observed heterozygosities (H_o) of all the seven loci were found to be significantly lower than those in the published database except for D1S243. Two loci namely, D3S1298 and D12S339, constantly showed deviation of HWE across the three sub-tribes (Table 2). F_{st} , F_{is} and F_{it} values computed are shown in Table 3.

Phenotypically, the Jahai are similar to the Habsyi population in Africa, tribes in the Andaman

Island and to the Aeta in the Philippines (<http://www.jakoa.gov.my/>). This nomadic population used to migrate from one place to another. Basically, their language (typically defined as Austroasiatic family) referred to by the same name, is a member of the Northern Aslian subgroup of the Aslian language, a branch of the Mon-Khmer language family. It is estimated that there are about 1,600 people in this tribe currently (statistics obtained from JAKOA, 2008).

Numbering 1,160 people in the tribe currently (statistics obtained from JAKOA, 2008), Bateqs are nomadic hunters and gatherers, linguistically belonging to the Mon-Khmer branch of the Austroasiatic language family, and they

Table 2. Allele frequencies, heterozygosity, and Hardy Weinberg Equilibrium (HWE) tests of three Negrito sub-tribes from Northern Peninsular Malaysia

Population	Locus	Observed no. allele (No. individuals)	HWE (Probability)	H _o (H _e)
Jahai (30)	D1S243	7	44.9271* (0.7411)	0.8400 (0.7856)
	D3S1298	6	51.9653 (0.0002)	0.2000 (0.7756)
	D7S517	6	12.6518* (0.9203)	0.6333 (0.7561)
	D8S1706	7	52.7500* (0.1994)	0.3846 (0.5710)
	D12S339	8	47.5536 (0.0120)	0.1250 (0.7795)
	D13S153	7	21.8889* (0.1108)	0.2857 (0.7551)
	D19S220	11	61.3946 (0.0000)	0.2963 (0.8395)
	Std Dev			0.2656 (0.0480)
Bateq (18)	D1S243	10	14.5518* (0.8448)	0.8889 (0.8642)
	D3S1298	7	48.9611 (0.0000)	0.4706 (0.7837)
	D7S517	7	12.2710* (0.6584)	0.7857 (0.7704)
	D8S1706	10	48.0273 (0.0007)	0.9091 (0.8140)
	D12S339	8	139.5030 (0.0000)	0.4167 (0.8194)
	D13S153	6	101.7420 (0.0000)	0.6250 (0.7656)
	D19S220	6	195.6306 (0.0000)	0.2000 (0.7133)
	Std Dev			0.2541 (0.0845)
Mendriq (14)	D1S243	8	27.5472* (0.4887)	0.8333 (0.8194)
	D3S1298	7	80.5750 (0.0000)	0.2857 (0.8036)
	D7S517	6	24.8000* (0.0527)	0.5833 (0.7292)
	D8S1706	9	50.9394* (0.0506)	0.5385 (0.7663)
	D12S339	7	60.8333 (0.0000)	0.4167 (0.8160)
	D13S153	5	27.7800 (0.0020)	0.4545 (0.6983)
	D19S220	7	25.8889* (0.2107)	0.4000 (0.7250)
	Std Dev			0.1753 (0.0490)

Expected heterozygosity values were computed using Nei's (1973)
Std Dev = standard deviation

*Do not deviate from Hardy Weinberg Equilibrium

phenotypically look similar to the Jahai. They are scattered around Northern Peninsular Malaysia and the Malaysian National Park (*Taman Negara*).

Mendriq is indeed an interesting population. They are phenotypically similar to the populations in the Andaman Island, the Philippines and Southern Thai; smaller in size but tougher. This population has existed since the Paleolithic age and speak a pure Mon-Khmer language. The number of people in the Mendriq population is dropping dramatically and currently there are only 288 people in this tribe as recorded by JKOA (2008). The actual reason for this depletion in population size is unknown, but it could be attributed to their lifestyle during earlier age, whereby the females may have taken traditional herbs to control the birth rate so as to ease their migration. In addition, the Mendriq has not been studied genetically in any serious way for over the last few decades.

The total number of alleles observed in Bateq, Jahai and Mendriq were 54, 52, 49 respectively, suggesting similar levels of polymorphisms among the populations. However, lower observed heterozygosity (H_o) values were obtained for these sub-tribes when compared to the published data (CEPH database; Tan *et al.*, 2002), indicate the effects of small population sizes (population

bottleneck). The values remained lower when we pooled the Negrito sub-tribes together. The lower value of H_o than the expected (H_e) suggested that inbreeding and genetic drift might have occurred in these Negrito sub-tribes. This is further supported by the F_{is} and F_{it} values in Table 3 (mean F_{is} = 0.3481; mean F_{it} = 0.3910). Furthermore, this finding was in line with earlier study on maternal lineage mtDNA (Hill *et al.*, 2006); while PCA analysis carried out by Peng and co-workers in 2010 proved that these populations faced severe genetic drift. The low value of F_{st} (mean 0.0658; Table 3) suggested high genetic similarity among the populations. Despite a small number of samples, this preliminary finding was expected in view of the small population size of Negritos (less than 3,000 individuals in total (<http://damak.jheoa.gov.my>) and high level of consanguineous marriage. Considering the stringent inclusion criteria of the subjects enrolled, we believe this finding is valid. However, a larger scale of investigation with larger sample size by including more tribes and more microsatellite markers should be done to confirm these results. Such study would help in our understanding of evolution and pre-historic human migration in the Malay Archipelago which until today remains largely a mystery. By understanding the evolution and migration, as well

Table 3. Summary of F-Statistics of seven loci studied among the Negrito sub-tribes

Locus	F _{is}	F _{it}	F _{st}
D1S243	-0.0377	0.0358	0.0708
D3S1298	0.5953	0.6275	0.0795
D7S517	0.1123	0.1649	0.0593
D8S1706	0.1484	0.2248	0.0898
D12S339	0.6032	0.6243	0.0533
D13S153	0.3848	0.4115	0.0434
D19S220	0.6065	0.6320	0.0649
Mean	0.3481	0.3910	0.0658

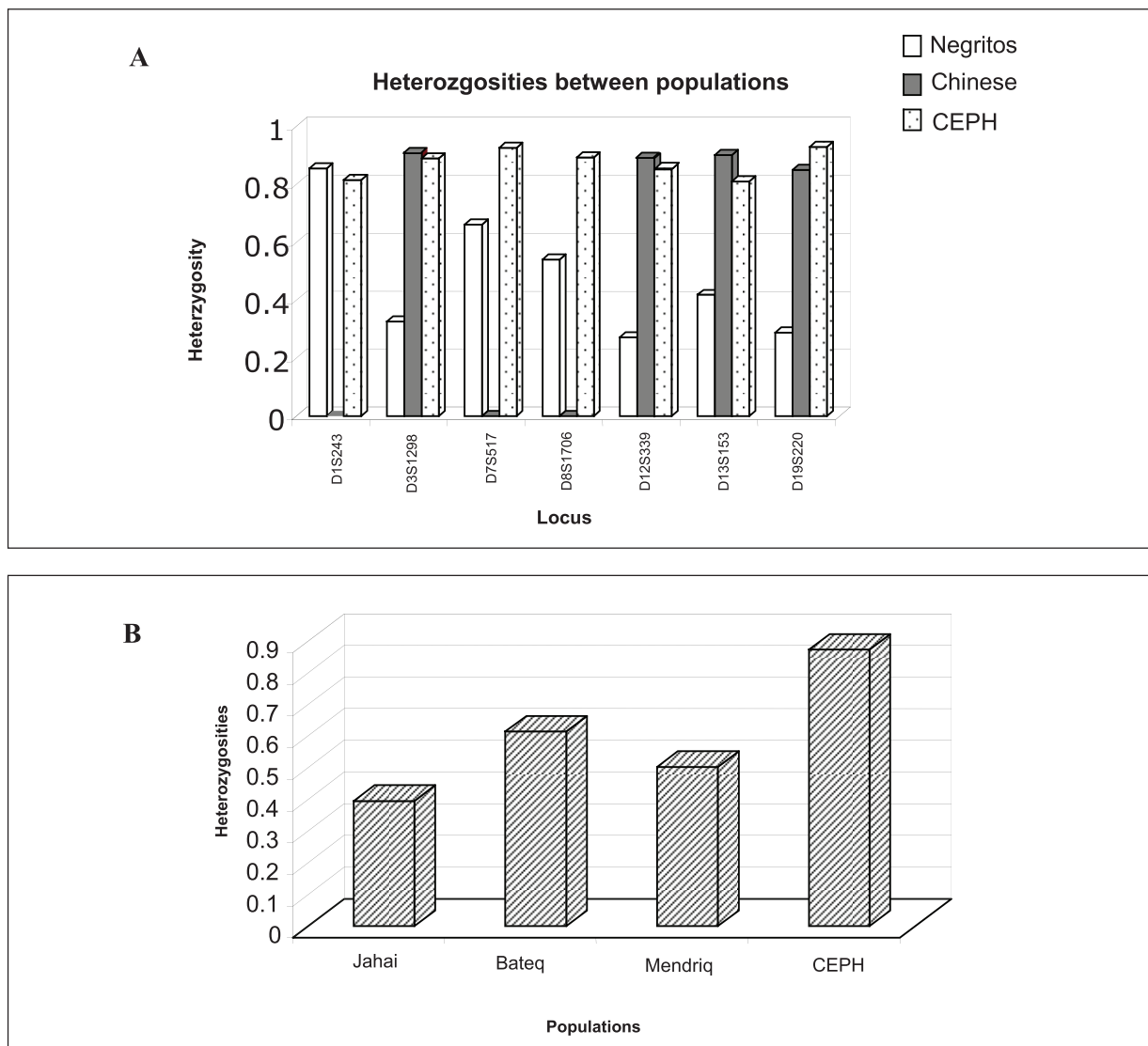


Fig. 3. (A) Comparison of mean heterozygosities of Negrito, Caucasian (<http://www.cephb.fr/cephdb/php/>) and Chinese populations (Tan *et al.*, 2002); (B) Heterozygosities of Jahai, Bateq, Mendriq and Caucasian

as genome diversity, it sheds lights to the various biomedical investigations of human diseases which may be prevalent in such populations and Southeast Asia in general.

ACKNOWLEDGEMENTS

This study was supported by Universiti Sains Malaysia (USM) Short Term Grant 304/PPSP/613508, Health Campus, Universiti Sains Malaysia. We thank the Department of Orang Asli Development Malaysia (*Jabatan Kemajuan Orang Asli, JAKOA*), for their support before and during the sampling expeditions; the staff nurses from Universiti Sains Malaysia for helping with sample collection; and the people who voluntarily and generously participating by providing their blood samples for this study.

REFERENCES

- Bellwood, P. 1993. Cultural and biological differentiation in Peninsular Malaysia: the last 10,000 years. *Asian Perspectives*, **32**: 37-60.
- Fernandes, A.T., Velosa, R., Jesus, J., Carracedo, A. & Brehm, A. 2003. Genetic differentiation of the Cabo Verde archipelago population analysed by STR polymorphisms. *Annals of Human Genetics*, **67**: 340-347
- Jin, L., Seielstad, M. & Xiao, C. 2001. *Genetic, linguistic and archeological perspectives on human diversity in Southeast Asia*. World Scientific Publishing. Singapore.
- Hill, C., Soares, P. & Mormina, M. 2006. Southeast Asians. *Molecular Biology and Evolution*, **23**: 2480-2491.
- Macaulay, V., Hill, C.M., Achilli, A., *et al.* 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondria genomes. *Science*, **308**: 1034-1036.
- Park, M.H., Kim, K.S., Lee, H.J., *et al.* 2008. Allelic frequencies and heterozygosities of microsatellite markers covering the whole genome in the Korean. *Journal Human Genetics*, **53**: 254-266.
- Payseur, B.A., Jing, P. & Haasl, R.J. 2010. A genomic portrait of human microsatellite variation, *Journal of Molecular Biology and Evolution*, doi: 10.1093/molbev/msq198.
- Peng, M-S., Quang, H.H., Dang, K.P., *et al.* 2010. Tracing the Austronesian footprint in mainland Southeast Asia: perspective from mitochondrial DNA, *Molecular Biology and Evolution*, **27**: 2417-2430.
- Tan, E., Wu, H., Yong, R., *et al.* 2002. Heterozygosities and allelic frequencies of a set of microsatellite markers used for genome-wide scans in a Chinese population. *Journal Human Genetics*, **47**: 623-631.
- Tan, S.G. 1979. Genetic relationship between Kadazans and fifteen other Southeast Asian races. *Pertanika*, **2**: 28-33.
- Tan, S.G. & Teng, Y.S. 1978. Saliva acid phosphotase and amylase in Senoi and Abridinal Malays and superoxide dismutase in various racial groups of Peninsular Malaysia. *Japan Journal of Human Genetics*, **23**: 133-138.
- Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., *et al.* 2009. The genetic structure and history of Africans and African Americans. *Science*, **324**: 1035-1044.