Chapter Four Discussion

4.1 Polymorphic site assays

Each genomic DNA extracted from blood samples of 210 healthy, unrelated Malaysian individuals (70 Chinese, 70 Indians, and 70 Malays) was used to assay for 5 polymorphic sites (*Hin*fI, Hump1, Hump2, O1, and O2). The results revealed that polymorphism was observed only at the *Hin*fI, Hump1, and Hump2 sites, whereas the O1 and O2 sites were monomorphic for these samples.

For *Hin*fI, Hump1, and Hump2 sites, each appeared to be at Hardy-Weinberg equilibrium, which is similar to the findings of Koh *et al.* (1993, 1994), and Monckton *et al.* (1993).

Tables 32 and 32a show that the distributions of *Hin*fI + and -, Hump1 C and G, and Hump2 C and T alleles among these 3 races were similar (not significant). Hence, the data indicated that these 3 races were closely related genetically.

Comparisons of the data among the Malaysian (Chinese, Indian, and Malay), Caucasian, and Japanese population samples (Tables 32 and 32a) revealed that the distributions of *Hin*fI + and -, and Hump1 C and G alleles among these samples were similar (not significant), except for the Hump1 alleles in the Malays/Caucasians. However, for the Hump2 C and T alleles among these samples, only the Chinese and Japanese have similar distributions, whereas the others were not. Thus, these comparisons implied that of the 3 Malaysian racial groups, the Chinese have the closest genetic relationship with the Japanese; on the other hand, the Malays have the most distant genetic relationship with the Caucasians.

Based on the data of *Hin*fI, Hump1, and Hump2 assays, approximately 67% (i.e., 140/210 x 100%) of the Malaysian individuals were heterozygous at one or more of the flanking polymorphic sites, compared to 71% of the Caucasians and 60% of the Japanese individuals (Monckton *et al.*, 1993), which can therefore have single alleles mapped by allele-specific MVR-PCR.

O1 and O2 polymorphic sites were not observed in the Malaysian tested. However, O1 polymorphic site was reported in the Africans and few Caucasians, in which the O1C-linked alleles have reduced mutation rate and allelic diversity. This O1C variant provides strong, though not conclusive, evidence for the mutation initiator model (Jeffreys *et al.*, 1994,1995; Monckton *et al.*, 1994). Since the O1 polymorphic site was not seen in the Malaysian population, it is expected that the MS32 in the Malaysians will have extreme allelic diversity similar to those observed in the Caucasians and the Japanese (Jeffreys *et al.*, 1991; Monckton *et al.*, 1993; Tamaki *et al.*, 1993).

Table 32 : Comparison of the allele frequencies at the MS32 flanking polymorphisms among the Malaysian, Caucasian, and Japanese population samples.

Locus	Allele	Malaysians						Caucasians		Japanese	
		Chinese		Indians		Malays					
		Freq.	No.	Freq.	No.	Freq.	No.	Freq.	No.	Freq.	No.
*Hin*fI	+	0.786	110	0.786	110	0.764	107	0.81	129	0.81	163
	-	0.214	30	0.214	30	0.236	33	0.19	31	0.19	39
Hump1	G	0.779	109	0.757	106	0.821	115	0.69	55	0.80	159
	C	0.221	31	0.243	34	0.179	25	0.31	25	0.20	41
Hump2	C	0.843	118	0.771	108	0.779	109	0.59	94	0.91	184
	T	0.157	22	0.229	32	0.221	31	0.41	66	0.09	18

Note:

The data of allele frequencies for the Caucasians and Japanese were obtained from Monckton (1992).

Freq. = Frequency

No. = Number

Table 32a : Pairwise comparisons by the heterogeneity G-test of the distributions of *Hin*fI + and -, Hump1 C and G, and Hump2 C and T alleles between different population samples.

Pairwise comparisons	*Hin*fI (+ and -)			Hump1 (C and G)			Hump2 (C and T)		
	G_H	df	P	G_H	df	P	G_H	df	P
Chinese vs Indians	0.000	1	P>0.995	0.180	1	0.70>P>0.50	2.305	1	0.20>P>0.10
Chinese vs Malays	0.184	1	0.70>P>0.50	0.805	1	0.50>P>0.30	1.893	1	0.20>P>0.10
Chinese vs Caucasians	0.194	1	0.70>P>0.50	2.188	1	0.20>P>0.10	24.411	1	0.001>P*
Chinese vs Japanese	0.230	1	0.70>P>0.50	0.133	1	0.75>P>0.70	3.639	1	0.10>P>0.05
Indians vs Malays	0.184	1	0.70>P>0.50	1.745	1	0.20>P>0.10	0.022	1	0.90>P>0.80
Indians vs Caucasians	0.194	1	0.70>P>0.50	1.242	1	0.30>P>0.25	11.681	1	0.001>P*
Indians vs Japanese	0.230	1	0.70>P>0.50	0.682	1	0.50>P>0.30	12.681	1	0.001>P*
Malays vs Caucasians	0.782	1	0.50>P>0.30	5.067	1	0.025>P>0.02*	12.693	1	0.001>P*
Malays vs Japanese	0.898	1	0.50>P>0.30	0.370	1	0.70>P>0.50	11.602	1	0.001>P*
Caucasians vs Japanese	0.000	1	P>0.995	3.534	1	0.10>P>0.05	53.944	1	0.001>P*

Note :

Appendix G shows the method of calculation for the results (Sokal and Rahlf, 1981).

G_H = χ^2 of the heterogeneity G-test

df = Degree of freedom

P = Probability

* = Significant in the distributions of alleles between different population samples

4.2 Haplotype assays

The statistical analyses for all haplotypes of three polymorphic sites (Hump1, *Hin*fI, and Hump2) in all three ethnic groups showed that a significant association among these three sites existed in the Malaysian haplotypes. This finding agrees with the data published by Monckton *et al.* (1993), in which the similar phenomena occurred in the Caucasian and Japanese haplotypes.

Significant association was also observed to exist between each pair of the polymorphic sites (*Hin*fI-Hump1, *Hin*fI-Hump2, and Hump1-Hump2) when calculations were made by using the 2 x 2 contingency tables.

Comparison of the haplotype frequencies at the MS32 flanking polymorphisms among the Malaysian (Chinese, Indian, and Malay), Caucasian, and Japanese population samples (Tables 33 and 33a) revealed that the haplotype G+C occurred at a higher frequency than the others in all samples. On the other hand, certain haplotypes did not exist in the population samples, for examples, the C-T in all populations; the C-C in the Malaysians and Caucasians; and the C+T in the Chinese and Japanese. This showed that the formations of a certain haplotypes were favoured over the others during the evolution of human species.

Pairwise comparisons of the distributions of 8 haplotypes between different population samples (Table 33b) showed that only the Chinese/Indians, Chinese/Malays, Chinese/Japanese, and Indians/Malays have the same distributions of haplotypes each, whereas the others were not. These results confirmed the findings in the polymorphic site assays (section 4.1), in which the Chinese, Indians, and Malays were closely related to each other genetically. Besides that, these results actually indicated that only the Chinese have a close genetic relationship with the Japanese, whereas the other races were not. This finding agrees with the data published by Bowcock *et al.* (1994), Hammer and Horai (1995), and Horai *et al.* (1996), in which the mainland Japanese shared great degree of genetic affinity with the Chinese from China. Since majority of the Malaysian Chinese

migrated from China, the genetic affinity between the Malaysian Chinese and mainland Japanese remained.

Table 33 : Comparison of the haplotype frequencies at the MS32 flanking polymorphisms among the Malaysian, Caucasian, and Japanese population samples.

Haplotype	Malaysians			Caucasians			Japanese		
H1-Hf-H2	Obsd.	f	Expd.	Obsd.	f	Expd.	Obsd.	f	Expd.
C-T	0	0.000	4.012	0	0.00	4	0	0.000	1
C-C	0	0.000	15.851	0	0.00	6	1	0.005	7
G-T	44	0.105	14.737	23	0.14	9	11	0.055	3
G-C	49	0.117	58.219	8	0.05	12	26	0.130	28
C+T	11	0.026	14.143	30	0.19	16	0	0.000	3
C+C	79	0.188	55.875	12	0.08	24	40	0.200	29
G+T	30	0.071	51.947	13	0.08	37	7	0.035	12
G+C	207	0.493	205.216	74	0.46	53	115	0.575	118

Note:

The data of haplotype frequencies for the Caucasians and Japanese were obtained from Monckton *et al.* (1993).

Obsd. = Observed number

f = Observed frequency

Expd. = Expected number

Table 33a : Comparison of the haplotype frequencies at the MS32 flanking polymorphisms among the Malaysian Chinese, Malaysian Indian, Malaysian Malay, Caucasian, and Japanese samples.

Haplotype	Malaysians						Caucasians		Japanese	
H1-Hf-H2	Chinese		Indians		Malays					
	Obsd	f	Obsd	f	Obsd	f	Obsd	f	Obsd	f
C-T	0	0.000	0	0.000	0	0.000	0	0.00	0	0.000
C-C	0	0.000	0	0.000	0	0.000	0	0.00	1	0.005
G-T	11	0.079	14	0.100	19	0.136	23	0.14	11	0.055
G-C	19	0.136	16	0.114	14	0.100	8	0.05	26	0.130
C+T	0	0.000	8	0.057	3	0.021	30	0.19	0	0.000
C+C	31	0.221	26	0.186	22	0.157	12	0.08	40	0.200
G+T	11	0.079	10	0.071	9	0.064	13	0.08	7	0.035
G+C	68	0.486	66	0.471	73	0.521	74	0.46	115	0.575

Note:

The data of haplotype frequencies for the Caucasians and Japanese were obtained from Monckton *et al.* (1993).

Obsd = Observed number

f = Observed frequency

Table 33b : Pairwise comparisons by the heterogeneity G-test of the distributions of 8 haplotypes between different population samples.

Pairwise comparisons	G_H	df	P
Chinese vs Indians	12.224	7	0.10<P<0.05
Chinese vs Malays	8.990	7	0.30<P<0.25
Chinese vs Caucasians	58.310	7	0.001<P*
Chinese vs Japanese	6.084	7	0.70<P<0.50
Indians vs Malays	3.990	7	0.80<P<0.75
Indians vs Caucasians	23.292	7	0.005<P<0.001*
Indians vs Japanese	21.556	7	0.005<P<0.001*
Malays vs Caucasians	30.068	7	0.001<P*
Malays vs Japanese	15.732	7	0.05<P<0.025*
Caucasians vs Japanese	79.582	7	0.001<P*

Note:

Appendix H shows the method of calculation for the results (Sokal and Rahlf, 1981).

G_H = χ^2 of the heterogeneity G-test

df = Degree of freedom

P = Probability

* = Significant in the distributions of haplotypes between different population samles

4.3 Direct DNA sequencing

In this study, a direct DNA sequencing method with an allele-specific primer 32-H1C and universal primer 32-O was used to sequence the 5' flanking region. A total length of 244 bases was obtained by this approach, which was fast, accurate and no cloning was required. This method is different from the previous studies (Armour *et al.*, 1989; Monckton *et al.*, 1993), in which the allele-specific primer 32-H1C was used to amplify the DNA strand with $H1^C$, and then separated from the DNA strand with $H1^G$ in homologous chromosome.

From the 12 samples tested, no new polymorphic sites were detected, while the existed polymorphic sites (Hump1, *Hin*fI, and Hump2) remained the same as those published in Monckton *et al.* (1993). No new polymorphic site was obtained in this study and this might due to the small amount of samples sequenced. Besides that, since only the DNA strand with $H1^C$ was sequenced, those new polymorphic sites in the DNA strand with $H1^G$ were definitely missed out from being sequenced. Thus, more samples are required to be sequenced, and allele-specific primer 32-H1G should be used besides primer 32-H1C, in order to enhance the chances of finding new polymorphic sites.

4.4 Conclusion

Five polymorphic sites (O2, *Hin*fI, Hump1, O1, and Hump2) at the DNA flanking the ultravariable end of the MS32 were assayed in 210 healthy, unrelated Malaysian individuals (70 Chinese, 70 Indians, and 70 Malays). Of the 5 sites, only the *Hin*fI, Hump1, and Hump2 were polymorphic, whereas the O1 and O2 were not. For these 3 polymorphic sites, each appeared to be at Hardy-Weinberg equilibrium. Subsequently, based on the data of polymorphic site assays, approximately 67% of the Malaysian individuals can have their single alleles mapped by allele-specific MVR-PCR.

In the haplotype assays, significant association was observed among the *Hin*fI, Hump1, and Hump2 sites. Significant association was also found in each pair of the polymorphic sites (*Hin*fI-Hump1, *Hin*fI-Hump2, and Hump1-Hump2). Besides that, in the Malaysian haplotypes, G+C occurred at the highest frequency, whereas C-T and C-C did not even exist. Meanwhile, haplotype C+T was also not found in the Malaysian Chinese samples. From the pairwise comparisons of the distributions of 8 haplotypes between different population samples, the results showed that the Malaysian Chinese, Indians, and Malays were closely related to each other genetically. The results also revealed that the Malaysian Chinese shared great genetic affinity with the Japanese.

In the DNA sequencing, a total length of 244 bases of the DNA flanking the ultravariable end of the MS32 was sequenced in 12 samples. However, no new polymorphic sites were detected, while the existing polymorphic sites remained the same.