ACE 2212
INVC ACE 2212
MSA - 24/12/97

ANALYSES OF THE DNA FLANKING THE HYPERVARIABLE MINISATELLITE MS32 (LOCUS D1S8) IN THE MALAYSIAN POPULATION

LEONG YIM KHENG

A dissertation submitted to the
Institute of Advanced Studies, University of Malaya,
in partial fulfilment of the requirements
for the Degree of Master of Biotechnology

March 1997

Perpustakaan Universiti Malaya
A507467478

Institute of Advanced Studies
University of Malaya
Kuala Lumpur
Malaysia

Abstract

Analyses of the DNA flanking the hypervariable minisatellite MS32 (locus D1S8) were carried out on 210 healthy, unrelated Malaysian individuals (70 Chinese, 70 Indians, and 70 Malays). Genomic DNA from each blood sample was extracted by using the rapid mini-preparation of genomic DNA protocol.

Five polymorphic sites (O2, *Hin*fI, Hump1, O1, and Hump2) were assayed for each DNA sample by using the PCR-based tests, restriction enzyme digestions, and gel electrophoresis. Of the 5 sites, only the *Hin*fI, Hump1, and Hump2 were polymorphic, whereas the O1 and O2 were monomorphic. For the *Hin*fI, Hump1 and Hump2 sites, each appeared to be at Hardy-Weinberg equilibrium, and their allele frequencies, power of discrimination (Pd), and heterozygosity (h) were calculated. Based on the data of *Hin*fI, Hump1, and Hump2 assays, approximately 67% of the Malaysian individuals were heterozygous at one or more of the flanking polymorphic sites, which can therefore have their single alleles mapped by the allele-specific minisatellite variant repeat mapping by polymerase chain reaction (MVR-PCR).

Haplotype assays were tested out only on the double heterozygous (*Hin*fI-Hump1, *Hin*fI-Hump2, and Hump1-Hump2) individuals by using the PCR-based tests, restriction enzyme digestions, and gel electrophoresis. Statistical analyses for all haplotypes of *Hin*fI, Hump1, and Hump2 in the Malaysian and each ethnic group showed that a significant association existed among these three sites. Significant association was also observed between each pair of the polymorphic sites (*Hin*fI-Hump1, *Hin*fI-Hump2, and Hump1-Hump2). Haplotype G+C appeared to be at the highest frequency, whereas haplotypes C-T and C-C, C+T did not exist in all samples and the Chinese, respectively.

Pairwise comparisons of the distributions of 8 haplotypes between different population samples showed that the Chinese, Indians, and Malays were closely related to each other genetically. The results also indicated that the Chinese shared great degree of genetic affinity with the Japanese, whereas the other races were not.

Direct DNA sequencing method with an allele-specific primer 32-H1C and a

PERPUSTAKAAN UNIVERSITI MALAYA

universal primer 32-O was used to sequence the 5' flanking region in 12 samples. A total length of 244 bases was obtained by this approach, 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 the homologous chromosome. The known polymorphic sites (*Hin*fI, Hump1 and Hump2) remained the same and no new polymorphic sites were detected.

Acknowledgments

I want to thank my supervisors, Professor Dr. Koh Chong Lek and Associate Professor Dr. Sam Choon Kook for their guidance in thinking, reading, experimenting, and writing of science; valuable suggestions and gave me the opportunity to work in the wonderful world of molecular biology.

I also want to thank Miss Lim Moo Eng and Mr. Edmund Sim for collecting the blood samples and their useful assistance.

I would like to thank Hann Ling, Yeong, Liz, Yusmin, Huey Sin, Ten Hong, Igor, Anuar, Nesa, Lina, Ping Chin, Chua, Ai Lam and Choon Weng for their help and great friendship. Special thanks to Hann Ling and Yeong for their active discussion and intellectual stimulation which really widen my horizon. Thanks to the staffs of the Dept. of Genetics and Cellular Biology, and IPT, particularly Mr. Gopal, for offering their assistance to me.

I like to thank Katsumata sensei, Tamaki sensei, Yamamoto sensei, Uchihi sensei, Naganawa-san, Yoshimoto-san, Huang-san, Nozawa-san and Okada-san for their warm hospitallity and assistance while I was writing my thesis in Japan. Special thanks to Yamamoto sensei for his great help in computing and friendship.

I am grateful to my family members for being thoughtful and caring through out my study.

List Of Contents

			Page
Abstract			i
Acknowledgments			iii
List Of Contents			iv
Abbreviations			viii
List Of Figures			xi
List Of Tables			xii
Chapter One : Introduction			1
1.1	:	Types of repetitive elements	1
1.2	:	Tandemly repetitive DNA	1
1.2.1	:	Satellite DNA	1
1.2.2	:	Midisatellite DNA	1
1.2.3	:	Minisatellite DNA	2
1.2.4	:	Microsatellite DNA	2
1.3	:	Minisatellite DNA and DNA polymorphism	2
1.4	:	Hypervariable minisatellite MS32	4
1.5	:	Minisatellite variant repeat mapping by the polymerase chain reaction (MVR-PCR)	5
1.6	:	DNA flanking the minisatellite MS32	6
1.7	:	Objectives of this study	7
Chapter Two : Materials And Methods			11
2.1	:	Blood samples	11
2.2	:	Oligonucleotide primers	11
2.3	:	Enzymes	11
2.4	:	Reagents and chemicals	12
2.5	:	Stock solutions and buffers	13
2.5.1	:	11.1x PCR mixture (mix)	13
2.5.2	:	10x Kinase buffer	14

2.5.3	:	20x Base mix	14
2.5.4	:	4x PCR termination mix	14
2.5.5	:	Stop solution (Formamide-dye mix)	15
2.5.6	:	Amberlite MB1	15
2.5.7	:	5x Acrylamide/bisacrylamide (38:2) [30% acrylamide stock solution]	15
2.6	:	Sterilization	15
2.7	:	Rapid mini-preparation of genomic DNA from blood samples	16
2.8	:	Determination of DNA concentrations	17
2.9	:	Polymorphic site and haplotype assays	17
2.9.1	:	General protocols for polymorphic site and haplotype assays	18
2.9.2	:	O2 assay	19
2.9.3	:	*Hin*fI assay	20
2.9.4	:	Hump1 assay	20
2.9.5	:	O1 assay	21
2.9.6	:	Hump2 assay	22
2.9.7	:	*Hin*fI-Hump1 assay	23
2.9.8	:	*Hin*fI-Hump2 assay	24
2.9.9	:	Hump1-Hump2 assay	25
2.10	:	Electrophoresis of DNA	26
2.11	:	Protocol for direct DNA sequencing	27
2.11.1	:	Preparation of double-stranded DNA (dsDNA)	27
2.11.2	:	Preparation of single-stranded DNA (ssDNA) or 1/2 PCR	28
2.11.3	:	Cycle sequencing	29
2.11.3.1	:	End labelling of primer	29
2.11.3.2	:	Sequencing reaction	29
2.11.4	:	Sequencing gel	29

2.11.4.1	:	Assembly of glass plates	30
2.11.4.2	:	Preparation of a 6% polyacrylamide sequencing gel	30
2.11.4.3	:	Polyacrylamide gel electrophoresis (PAGE) and autoradiography	31
2.12	:	Statistical analyses	33
Chapter Three : Results			35
3.1	:	Assays for polymorphic sites	35
3.1.1	:	PCR assays with or without restriction enzyme digestions	35
3.1.2	:	Statistical analyses of results from polymorphic site assays	42
3.2	:	Haplotype assays	46
3.2.1	:	PCR assays with or without restriction enzyme digestions	46
3.2.2	:	Statistical analyses of results from haplotype assays	50
3.3	:	Direct DNA sequencing	58
Chapter Four : Discussion			61
4.1	:	Polymorphic site assays	61
4.2	:	Haplotype assays	65
4.3	:	Direct DNA sequencing	70
4.4	:	Conclusion	71
References			72
Appendixes :			80
Appendix A	:	Results of MS32 (Locus D1S8) typing	80
Appendix B	:	Methods of calculation for polymorphic site assays	91
Appendix C	:	(i) Results of *Hin*fI-Hump1 assay	93
		(ii) Results of *Hin*fI-Hump2 assay	93
		(iii) Results of Hump1-Hump2 assay	95
Appendix D	:	Method of calculation for haplotype assays	97
Appendix E	:	Method of calculation for each pair of polymorphic sites	98
Appendix F	:	Samples for DNA sequencing	99

Appendix G : 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 100

Appendix H : Pairwise comparisons by the heterogeneity G-test of the distributions of 8 haplotypes between different population samples 110

Abbreviations

Most of the abbreviations used are standards. However, attention is drawn to the following:

%	:	percentage
(v/v)	:	volume per volume
(w/v)	:	weight per volume
<	:	less than
>	:	more than
A	:	adenine
APS	:	ammonium persulphate
ATP	:	adenosine 5'-triphosphate
bp	:	base pair
BPB	:	bromophenol blue
BSA	:	bovine serum albumin
C	:	cytosine
Ci	:	Curie
cm	:	centimetre
dATP	:	deoxyadenosine 5'-triphosphate
dCTP	:	deoxycytidine 5'-triphosphate
ddNTP	:	dideoxyribonucleoside 5'-triphosphate
df	:	degree of freedom
dGTP	:	deoxyguanosine 5'-triphosphate
dH_2O	:	distilled water
DNA	:	deoxyribonucleic acid
dNTP	:	deoxyribonucleoside 5'-triphosphate
dsDNA	:	double-stranded DNA
dTTP	:	deoxythymidine 5'-triphosphate
EDTA	:	ethylenediaminetetraacetate
EtBr	:	ethidium bromide
EtOH	:	ethanol
Fig.	:	figure
g	:	gramme
G	:	guanine
h	:	heterozygosity
H1	:	Hump1
H2	:	Hump2

PERPUSTAKAAN UNIVERSITI MALAYA

HCl	:	hydrochloric acid
Hf	:	*Hin*fI
hr	:	hour
HVR	:	hypervariable region
kb	:	kilobase pairs
LINE	:	long interspersed repeat element
LTR	:	long terminal repeat
M	:	molar
mA	:	milliampere
Mb	:	megabase pairs
mCi	:	milliCurie
mg	:	milligramme
$MgCl_2$	:	magnesium chloride
min	:	minute
ml	:	millilitre
mm	:	millimetre
mM	:	millimolar
mMol	:	millimole
MVR-PCR	:	minisatellite variant repeat mapping by the polymerase chain reaction
NaCl	:	sodium chloride
NaOAc	:	sodium acetate
NH_4OAc	:	ammonium acetate
OD	:	optical density
P	:	probability
PAGE	:	polyacrylamide gel electrophoresis
Pd	:	power of discrimination
psi	:	pound per square inch
RFLP	:	restriction fragment length polymorphism
rpm	:	revolutions per minute
RTVL-I	:	retrovirus-like element
sdH_2O	:	sterile/distilled water
SDS	:	sodium dodecyl sulphate
sec	:	second
SINE	:	short interspersed repeat element
SSC	:	sodium chloride-sodium citrate
ssDNA	:	single-stranded DNA

T	:	thymine
TBE	:	Tris-borate EDTA
TEMED	:	N, N, N', N'- tetramethylenediamine
Tris	:	Tris (hydroxymethyl) methylamine
U	:	unit
uv	:	ultraviolet
VNTRs	:	variable number of tandem repeats
°C	:	degree Celsius
γ-^{33}P	:	radioisotope gamma phosphorus-33
ηg	:	nanogramme
ηm	:	nanometre
µg	:	microgramme
µl	:	microlitre
µm	:	micrometre
µM	:	micromolar
χ^2	:	Chi square

List Of Figures

Page

Fig. 1 : Dispersed repeat elements flanking the human minisatellite MS32. 9

Fig. 2 : Polymorphisms in the 5' flanking region of the minisatellite MS32. 10

Fig. 3 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from *Hin*fI assay. 37

Fig. 4 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from O2 assay. 38

Fig. 5 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from Hump1 assay. 39

Fig. 6 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from O1 assay. 40

Fig. 7 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from Hump2 assay. 41

Fig. 8 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from *Hin*fI-Hump1 assay. 47

Fig. 9 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from *Hin*fI-Hump2 assay. 48

Fig. 10 : Ethidium bromide-stained 4% (w/v) agarose gel (NuSieve GTG: SeaKem LE, 3:1) of DNA bands from Hump1-Hump2 assay. 49

Fig. 11 : Autoradiogram of DNA sequences sequenced by hot primer 32-O. 59

Fig. 12 : DNA sequence between primers 32-H1C and 32-O. 60

List Of Tables

		Page
Table 1 :	Primers.	11
Table 2 :	Restriction endonucleases and buffers.	12
Table 3 :	Reagents for 11.1x PCR mix.	13
Table 4 :	Reagents for 10x Kinase buffer.	14
Table 5 :	Reagents for 20x Base mix.	14
Table 6 :	Reagents for 4x PCR termination mix.	15
Table 7 :	Reagents for stop solution.	15
Table 8 :	PCR reagents for O2 assay.	19
Table 9 :	*Hae*III digestion for O2 assay.	20
Table 10 :	*Hin*fI digestion for *Hin*fI assay.	20
Table 11 :	PCR reagents for Hump1 assay.	21
Table 12 :	*Bsp*1286I digestion for Hump1 assay.	21
Table 13 :	PCR reagents for O1 assay.	22
Table 14 :	*Scr*FI digestion for O1 assay.	22
Table 15 :	PCR reagents for Hump2 assay.	23
Table 16 :	PCR reagents for *Hin*fI-Hump1 assay.	23
Table 17 :	*Hin*fI digestion for *Hin*fI-Hump1 assay.	24
Table 18 :	PCR reagents for *Hin*fI-Hump2 assay.	24
Table 19 :	*Hin*fI digestion for *Hin*fI-Hump2 assay.	25
Table 20 :	PCR reagents for Hump1-Hump2 assay.	25
Table 21 :	PCR reagents for dsDNA amplification.	27
Table 22 :	Reagents for 1/2 PCR.	28
Table 23 :	Reagents for sequencing reaction.	29
Table 24 :	Reagents for 6% polyacrylamide sequencing gel.	30
Table 25 :	Statistical analyses for *Hin*fI assay.	43
Table 26 :	Statistical analyses for Hump1 assay.	44

Table	27 :	Statistical analyses for Hump2 assay.	45
Table	28 :	Statistical analyses for Malaysian haplotypes.	51
Table	28a :	Statistical analyses for Malaysian Chinese haplotypes.	52
Table	28b :	Statistical analyses for Malaysian Indian haplotypes.	53
Table	28c :	Statistical analyses for Malaysian Malay haplotypes.	54
Table	29 :	2 x 2 contingency table for *Hin*fI and Hump1 polymorphic sites.	55
Table	30 :	2 x 2 contingency table for *Hin*fI and Hump2 polymorphic sites.	56
Table	31 :	2 x 2 contingency table for Hump1 and Hump2 polymorphic sites.	57
Table	32 :	Comparison of the allele frequencies at the MS32 flanking polymorphisms among the Malaysian, Caucasian, and Japanese population samples.	63
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.	64
Table	33 :	Comparison of the haplotype frequencies at the MS32 flanking polymorphisms among the Malaysian, Caucasian, and Japanese population samples.	67
Table	33a :	Comparison of the haplotype frequencies at the MS32 flanking polymorphisms among the Malaysian Chinese, Malaysian Indian, Malaysian Malay, Caucasian, and Japanese samples.	68
Table	33b :	Pairwise comparisons by the heterogeneity G-test of the distributions of 8 haplotypes between different population samples.	69