MUTATION SCREENING AND ANALYSIS OF THE *APC* GENE IN MALAYSIAN FAP PATIENTS

FIQRI DIZAR KHAIDIZAR

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2012

MUTATION SCREENING AND ANALYSIS OF THE *APC* GENE IN MALAYSIAN FAP PATIENTS

FIQRI DIZAR BIN KHAIDIZAR

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2012

Acknowledgement

I would first like to thank Professor Dr. Zulqarnain Mohamed for his supervision of this research project as well as his many ideas and constructive critisisms offered during the drafting of this thesis and Professor Datuk Dr. Rohana Yusof from the Faculty of Medicine, University of Malaya for allowing me to use the cell culture facility. Special thanks also go out to Dr. Sally Teh, Prof. Tan Peng Gwen and Dr. Sherry Bozo for their help in preparing this thesis and not forgetting to the members of the BGM and the Molecular Medicine Laboratory for their assistance and companionship throughout this whole wonderful journey.

And lastly, I would like to thank Khaidizar Kadir Mahidin and Wan Rokeah Meor Ghazali for their unconditional love and support.

ABSTRACT

Mutations in the *APC* tumour suppressor gene cause Familial Adenomatous Polyposis (FAP), a predisposition to colorectal cancer. This condition is characterized by adenomatous polyp formation in the colon and rectum which may advance to malignancy if not detected and treated early. FAP is 100% penetrant and segregates in an autosomal dominant manner. Screening for mutation in the *APC* gene is part of the preemptive genetic screening and risk assessment efforts for individuals with high risk of FAP. Knowledge of mutation spectrum that is prevalent in a certain population is essential to ensure precise prognosis of FAP by genetic testing. As such, this study was done to screen for pathogenic *APC* mutations in a set of Malaysian FAP patients and to reveal more information on the local *APC* mutation spectrum. DNA from each patient's blood sample were purified and subjected to PCR-SSCP analysis. Each exon of the *APC* gene was amplified by PCR using exon-specific primers. PCR-amplified products were then denatured and separated on MDE gel for approximately 16 hours. The gel was later stained by silver staining. Samples with distinct band mobility shift were sent for DNA sequencing to determine the nature of the sequence variation. Three truncating mutations and a single nucleotide polymorphism (SNP) were detected from the screening. Two of the mutations detected (c.847C>T and c.1690C>T) are point mutations that created premature stop codons while another mutation (c.875-876insT) is a frameshift mutation. c.847C>T and c.1690C>T mutations have been reported extensively in many cohorts while mutation c.875-876insT is believed to be novel. The SNP detected in this study (c.1635G>A), on the other hand, has been reported in SNP databases. c.847C>T, c.1690C>T and c.875-876insT are expected to cause truncation of the APC protein, which is consistent with other previously reported *APC* gene mutations. Truncated APC protein can only weakly maintain cytoplasmic β-catenin concentration at the baseline level. This condition promotes hyperactivation in the

transcription of Wnt target genes *en route* to colorectal carcinogenesis. The novel mutation detected here has expanded the knowledge of the *APC* mutations in the multiracial population of Malaysia. This would further assist in the establishment of a local *APC* mutation database that would benefit genetic testing and risk assessment strategies of FAP in the future.

A splicing assay using a minigene expression system was also carried out to investigate the effect of c.847C>T (p.R283X) mutation on exon splicing *in vitro* since the mutation was found to reside in an exonic splicing enhancer (ESE) motif. A minigene carrying the mutation was constructed using splicing by overlap extension–PCR (SOE-PCR) technique and cloned into a mammalian expression vector. The vector was then transfected into HepG2 cultured cells. mRNA was later extracted from the cells 48 hours after incubation and subjected to reverse transcription PCR (RT-PCR) for analysis. The mutation was expected to cause skipping of exon 8 from the mature minigene transcript. cDNA analysis showed that the expected transcript *sans* exon 8 was absent. Instead, five transcripts of variable length were observed. Sequencing of these variants revealed that they were the minigene's mRNA products of variable degree of splicing. The absence of the expected splicing product was thought to be due to the strength of the intron 7-exon 8 junction to define the boundary without necessitating the function of the adjacent ESE motif during splicing of exon 8, hence preserving exon 8 in the mature transcript.

ABSTRAK

Mutasi dalam gen Adenomatous Polyposis Coli (*APC*) menyebabkan penyakit Familial Adenomatous Polyposis (FAP) yang merupakan penyakit yang boleh membawa kepada pembentukan barah usus besar. Penyakit ini dapat dikenalpasti melalui pembentukan adenoma berbentuk polip di mana-mana kawasan sepanjang dinding usus besar dan rektum. Kegagalan dalam mengesan dan merawat pembentukan polip di peringkat awal akan meningkatkan risiko pembentukan barah usus di lewat usia. Penyakit FAP adalah penyakit yang boleh diwarisi secara autosom dominan dengan kadar 'penetrance' 100%. Pengenalpastian mutasi dalam gen *APC* adalah salah satu cara dalam strategi pengurusan risiko bagi individu berisiko tinggi menghidap penyakit FAP. Pengetahuan tentang jenis-jenis mutasi yang terdapat dalam gen *APC* bagi populasi masyarakat di Malaysia adalah sangat penting bagi memastikan prognosis terhadap penyakit FAP secara ujian genetik dapat dilakukan dengan tepat. Sehubungan dengan itu, kajian ini telah dijalankan bagi mengenalpasti jenis mutasi gen *APC* yang terdapat dalam pesakit-pesakit FAP di Malaysia. Asid deoksiribonukleik (DNA) daripada sel-sel darah putih pesakit FAP yang dikaji telah diasingkan dan dianalisa menggunakan teknik PCR-SSCP. Setiap salinan 'exon' dalam gen *APC* telah digandakan bilangannya secara PCR sebelum salinan-salinan DNA ini dipisahkan melalu teknik SSCP menggunakan gel MDE sebagai media pemisah selama 16 jam. Gel tersebut telah diwarnakan dengan menggunakan larutan perak nitrat. Pemerhatian dibuat bagi mengenalpasti kehadiran bentuk jaluran DNA yang berbeza apabila perbandingan dibuat antara sampel DNA pesakit dan sampel DNA bukan pesakit. Sampel yang didapati berbeza jaluran daripada sampel bukan pesakit akan dihantar untuk proses penjujukan nukleotida bagi mengenalpasti jenis mutasi yang terdapat di dalam sampel tersebut. Daripada analisis, kajian telah menjumpai dua mutasi titik (point mutation), setiap satu di dalam 'exon' 8 dan 13. Satu mutasi selitan (insertion) telah dijumpai juga

di dalam ‘exon’ 8 manakala satu nukleotida ‘polymorphic’ juga telah dikenalpasti. Dua mutasi titik tersebut (c.847C>T dan c.1690C>T) telahpun dilaporkan dalam penerbitan sebelum ini. Namun, didapati tiada penerbitan yang melaporkan tentang mutasi selitan (c.875-876insT) yang telah dijumpai dalam kajian ini. Ini menunjukkan bahawa mutasi ini mungkin merupakan mutasi baru.

Satu ujikaji kesan mutasi c.847C>T ke atas proses penyambatan ‘exon’ 8 telah dijalankan melalui kaedah pengekspresan gen *APC* mini di dalam kultur sel hepatosit mamalia secara *in vitro*. Bahagian gen *APC* yang mengandungi ‘exon’ 7, 8 dan 9 telah diklon ke dalam ‘plasmid vector’ pengekspres khusus bagi sel mamalia dan telah dipindahmasuk ke kultur sel hepatosit mamalia. Selepas 48 jam, mRNA yang dihasilkan daripada gen *APC* mini ini telah diasingkan dan dimurnikan sebelum bilangan salinan mRNA tersebut digandakan melalui kaedah ‘reverse transcriptase-PCR’ menggunakan oligonukleotida atau ‘primer’ yang spesifik kepada gen *APC* mini tersebut. Hasil daripada ujikaji ini menunjukkan bahawa produk pengekspresan gen *APC* mini tersebut hadir dalam pelbagai saiz. Kesimpulan yang dapat dibuat daripada pemerhatian ini adalah kehadiran mutasi c.847C>T tidak memberi kesan terhadap penyambatan ‘exon’ 8 dan pada masa yang sama proses penyambatan ‘exon’ di dalam kultur sel hepatosit mamalia tidak menghasilkan produk penyambatan yang spesifik.

Contents

Page

Abstract / Abstrak **ii**

Acknowledgement **vi**

List of figures **vii**

List of tables **ix**

Abbreviations **x**

Chapter 1: Introduction 1

Chapter 2: Literature Review

2.1: Cancer 4

2.1.1: Cancer definition 4

2.1.2: Carcinogenesis 5

2.1.3: Types of cancer 6

2.1.4: Anatomy of colon and rectum 8

2.2: Colorectal cancer 9

2.2.1: Epidemiology 9

2.2.2: Aetiology and risk 10

2.2.3: Diagnosis and treatments 10

2.3: Familial Adenomatous Polyposis (FAP) 11

2.4: *Adenomatous Polyposis Coli (APC)* gene 12

2.4.1: *APC* gene discovery 12

2.4.2: Background of *APC* gene 13

2.4.3: APC protein functions 15

2.5: Mutation screening 17

2.5.1: Methods for mutation screening 17

2.5.2: Single-strand Conformation Polymorphism (SSCP) 17

2.5.2.1: Critical parameters of SSCP 18

2.6: Objectives of study 19

Chapter 3: Methodology

3.1: Study subjects and samples 21

3.2: Genomic DNA extraction 22

3.3: DNA quantitation 23

3.4: Gel electrophoresis 24

3.4.1: Gel preparation 24

3.4.2: Electrophoresis 25

3.5: Polymerase Chain Reaction (PCR) 25

3.5.1: Primer design 25

3.5.2: PCR mastermix preparation 26

3.5.3: Amplification by PCR 26

3.6: Single strand conformation polymorphism (SSCP) analysis 27

3.6.1: Polyacrylamide Gel Electrophoresis (PAGE) 27

3.6.1.1: Gel cast assembly 27

3.6.1.2: Gel casting 28

3.6.1.3: Sample preparation 29

3.6.1.4: SSCP gel electrophoresis 29

3.6.2: Gel staining 30

3.7: PCR product clean-up 31

3.7.1: Direct DNA purification using QIAquick PCR purification kit 31

3.7.2: DNA purification using QIAquick gel extraction kit 32

3.8: DNA sequencing 33

3.9: Splicing assay 33

3.9.1: Amplification of exon 7, exon 8 and exon 9 by PCR 34

3.9.2: DNA cloning in bacterial hosts *Escherichia coli* 34

3.9.2.1: Preparation of bacterial growth broth and plated media 34

3.9.2.2: Preparation of competent cells 35

3.9.2.3: pGEM-T Easy cloning vector-based DNA cloning 36

(a) Ligation 36

(b) Transformation 37

(c) Colony selection by PCR 37

3.9.3: Preparation of Plasmid DNA by alkaline lysis with SDS 38

3.9.4: Site-directed mutagenesis 40

3.9.4.1: Mutant strand synthesis reaction preparation 41

3.9.4.2: Thermal cycling parameter 42

3.9.4.3: Transformation of XL1-Blue supercompetent cells 42

3.9.4.4: Colony selection 43

3.9.5: Minigene construction using splicing by overlap extension PCR (SOE-PCR) 43

3.9.5.1: Primer design 43

3.9.5.2: PCR amplification of the exons using overlap primers 44

3.9.5.3: SOE-PCR 46

3.9.6: Cloning of minigene into pTARGET mammalian expression 47

vector

3.10.6.1: *Eco*R1 digestion 47

3.10.6.2: Ligation 47

3.10.6.3: Transformation 47

3.10.6.4: Colony selection 48

3.9.7: Cell culture 48

3.9.7.1: Preparation of culture medium 48

3.9.7.2: Cell maintenance 49

3.9.7.3: Cell cryopreservation and revival 50

3.9.7.4: Cell enumeration 51

3.9.8: Transfection 51

3.9.9: RNA extraction 52

3.9.10: Reverse Transcriptase PCR (RT-PCR) 53

3.9.10.1: RT-PCR mastermix preparation 53

3.9.10.2: Thermal cycling profile 54

Chapter 4: Results

4.1: PCR-SSCP assay 55

4.1.1: PCR amplification of *APC* gene exons 55

4.1.2: SSCP analysis 55

4.1.2.1: SSCP analysis of exon 8 55

4.1.2.2: SSCP analysis of exon 13 56

4.2: DNA sequencing 58

4.2.1: Sequencing result of exon 8 from sample 0X-1 58

4.2.2: Sequencing result of exon 8 from sample 0X-2 58

4.2.3: Sequencing result of exon 13 from sample 06-1 60

4.2.4: Sequencing result of exon 13 from samples 06-2 to 06-7 60

4.3: Splicing assay 61

4.3.1: Minigene construction 61

4.3.2: Site-directed mutagenesis 62

4.3.3: Restriction enzyme assay 64

4.4: RT-PCR 64

4.5: DNA sequencing and analysis 66

Chapter 5: Discussion

5.1: Mutation screening by Polymerase Chain Reaction - Single strand Conformation Polymorphism (PCR-SSCP) method 69

5.2: *APC* mutations 72

5.3: The genotype-phenotype correlation of extracolonical FAP features 75

5.4: c.1635G>A polymorphism 77

5.5: *In vitro* splicing assay 77

Chapter 6: Conclusion 80

References 81

Appendices 90

Appendix A Reagents and Solutions 90

Appendix B List of primers 92

Appendix C Genomic sequences of exon 8 and exon 13 of the *APC* gene. 94

Appendix D Query results from online *APC* mutation databases. 95

Appendix E Proceedings and awards 98

List of table

Page

Table 2.1 Classification of FAP severity. 15

Abbreviations

%	percentage
µl	microlitre
µM	micromolar
AFAP	attenuated familial adenomatous coli
Ala	alanine
APC	adenomatous polyposis coli
APS	ammonium persulfate
Arg	arginine
BMD	Becker muscular dystrophy
bp	base pair
BSA	bovine serum albumin
c.	cDNA
$CaCl_2$	calcium chloride
cDNA	complementary DNA
cm	centimetre
c-myc	cellular-myelocytomatosis viral oncogene homolog
CO_2	carbon dioxide
COSMIC	Catalogue of somatic mutation in cancer
CRC	colorectal cancer
DGGE	denaturing gradient gel electrophoresis
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
Dsh	disshelved

E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ESE	exonic splicing enhancer
et al.	*et alii* (and other people)
EtBr	ethidium bromide
FAP	familial adenomatous polyposis
FBS	fetal bovine serum
Fz	frizzled
g	gravity
g.	genomic
GC	guanine-cytosine
Glu	glutamic acid
GSK	glycogen synthase kinase
HDA	heteroduplex analysis
HDLG	human disc large
HRMA	high resolution melting analysis
Ile	isoleucine
Ins	insertion
IPTG	isopropyl β-D-1-thiogalactopyranoside
kb	kilobase
kDa	kilo Dalton
K-*ras*	Kirsten- rat sarcoma
LB	Luria Bertani
lb/sq	pounds per square foot
Leu	leucine
LOVD	Leiden open variation database

M	Molar
MCR	mutation cluster region
MDE	mutation detection enhancer
Met	methionine
$MgCl_2$	magnesium chloride
Min	murine intestinal neoplasia
ml	millilitre
mM	milli Molar
mm	millimetre
MMTV	mouse mammary tumour virus
mRNA	messenger RNA
NaOH	sodium hydroxide
NAS	nonsense-associated altered splicing
ng	nanogram
^oC	degree Celsius
OD	optical density
p.	protein
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pfu	*Pyrococcus furiosus*
Phe	phenylalanine
pmol	picomole
PTC	premature termination codon
PTT	protein truncation test
ras	rat sarcoma

RE	restriction enzyme
RFLP	Restriction fragment length polymorphism
RNA	ribonucleic acid
RNase	ribonuclease
rpm	rotation per minute
RT-PCR	reverse-transcriptase polymerase chain reaction
SDM	site-directed mutagenesis
SDM-F	site-directed mutagenesis-forward
SDM-R	site-directed mutagenesis-reverse
SDS	sodium dodecyl sulfate
Ser	serine
SF2/ASF	splicing factor 2/ alternative splicing factor
SNP	single nucleotide polymorphism
SOE-PCR	splicing by overlap extension-PCR
SRp40	serine rich protein 40
SRp55	serine rich protein 55
SSCP	single-strand conformation polymorphism
Taq	*Thermus aquaticus*
TBE	tris borate EDTA
Tcf/LEF	T-cell factor/lymphocyte enhancer factor
TEMED	tetramethyl-ethylenediamine
Thr	threonine
Tris	tris(hydroxymethyl)amino methane
U	unit
UMD	Universal mutation database
UTR	untranslated region

UV	ultraviolet
V	volt
v/v	volume per volume
Val	valine
w/v	weight per volume
Wnt	wingless-type
X-gal	bromo-chloro-indolyl-galactopyranoside
YAC	yeast artificial chromosome
β-cat	β-catenin