STUDY OF GUT MICROBIOME AND METABOLOMICS PROFILING IN COLORECTAL CANCER PATIENTS

LUCKY GOH POH WAH

THESIS SUBMITTED IN FULFILLMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

FACULTY OF SCIENCE AND NATURAL RESOURCES
UNIVERSITI MALAYSIA SABAH
2019

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL: **STUDY OF GUT MICROBIOME AND METABOLOMICS PROFILING IN COLORECTAL CANCER PATIENTS**

IJAZAH: **DOKTOR FALSAFAH (BIOTEKNOLOGI)**

Saya **LUCKY GOH POH WAH**, Sesi **2016-2019**, mengaku membenarkan tesis Doktoral ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

☐	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)
☑	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
☐	TIDAK TERHAD	

Disahkan Oleh,

LUCKY GOH POH WAH
DS1521009T

AZLYNNE MOHD. JOHAN @ JACKLYNE
PUSTAKAWAN
UNIVERSITI MALAYSIA SABAH
(Tandatangan Pustakawan)

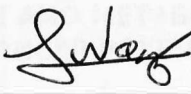
Tarikh : 12 September 2019

(Prof. Madya Dr. Lee Ping Chin)
Penyelia

UMS
UNIVERSITI MALAYSIA SABAH

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, equations, summaries and references, which have been duly acknowledged.



Lucky Goh Poh Wah
DS1521009T

CERTIFICATION

NAME : **LUCKY GOH POH WAH**

MATRIC NO. : **DS1521009T**

TITLE : **STUDY OF GUT MICROBIOME AND METABOLOMICS PROFILING IN COLORECTAL CANCER PATIENTS**

DEGREE : **DOCTOR OF PHILOSOPHY (BIOTECHNOLOGY)**

VIVA VOCE DATE : **1ST AUGUST 2019**

CERTIFIED BY

Signature

SUPERVISOR
Assoc. Prof. Dr. Lee Ping Chin

ACKNOWLEDGEMENTS

I would like to express my upmost gratitude to my supervisor associate professor Dr Lee Ping Chin for her unlimited guidance, support and patience given to me through the entire study. I would like to thanks to the surgeon at the Department of Surgery, Queen Elizabeth Hospital, Kota Kinabalu, Dr Edwin See Un Hean for his kind assistance in sample collection.

The research presented in this dissertation was performed at University Malaysia Sabah (UMS) and supported by Ministry of Education Malaysia, Fundamental Research Grant Scheme (FRGS0428-SG-1/2015) and Universiti Malaysia Sabah Postgraduate Research Grant (GUG0092-SG-2/2016). I also would like to thank the Ministry of Education Malaysia for awarding PhD scholarship under MyBrain15 to allow the completion of this doctoral degree.

Lastly, I am grateful to my lab colleagues for their willingness to share their knowledge and idea. Finally, I would like to express my sincere appreciation to my family for their continuous support and encouragement throughout this study.

Thank you.

Lucky Goh Poh Wah
12 March 2019

ABSTRACT

The human gastrointestinal tract harbors millions of gut microbiota. The gut microbiota has been linked to various cancers, including gastric, liver and colorectal cancers (CRC). Despite the profound knowledge and understanding of CRC development, CRC incidence rate is increasing every year and the trend is shifting to younger age every year. This signals the need to understand the role of gut microbiota and host interactions in CRC. Therefore, this study is aimed to investigate the gut microbiota and its association to annotated host metabolites that play a potential role in the CRC with the support from gene expression analysis. A total of 35 paired case-control intestinal mucosal tissue sample were collected from CRC patients and subjected to multiple analyses. Thirty case-control tissues were subjected to gut microbiota analysis using 16S rRNA next generation sequencing. Twenty-three case-control tissues were used in metabolomics analysis using mass spectrometry. Twenty-eight case-control tissues were subjected to gene expression analysis using real-time PCR. The composition of gut microbiota, alpha diversity metrics (chao1, goods coverage, observed OTUs, Shannon index, Simpson's index, and phylogenetic distance) and beta diversity metrics (ANOSIM, PERMANOVA, PERMDISP and MRPP) shows that the gut microbiota were stable in CRC patients despite the difference in tissue pathology (cancer *vs* normal). Genus level analysis identified *Campylobacter* from the *Proteobacteria* phylum was significantly (*p-value* < 0.05) more abundant in cases. Functional content profiling of the gut microbiota shows that glycosyltransferases, lipopolysaccharide biosynthesis proteins and its contents were also significantly (*p-value* < 0.05) higher in cases. Moreover, primary bile acid biosynthesis was most significantly lower (*p-value* < 0.05) in cases followed by bacterial toxins as well as pentose and glucuronate interconversions. Further analysis shows *Actinobacteria* and *Proteobacteria* phyla commonly contributed to the enzymes implicated in primary bile acid biosynthesis. Furthermore, metabolomics analysis annotated a total of nine known and two unknown polar metabolites as well as 19 glycerophospholipids and nine unknown non-polar metabolites with VIP score ≥ 1.5 and *p-value* < 0.05 that were elevated in cases. Inferences on the metabolites suggested the carnitine system that plays a part in beta-oxidation for energy production were involved in CRC which warrants further investigation using gene expression. Expression of *NR1H4/FXR* gene that was involved in bile acid signaling were significantly reduced (3.32-$\log_2$ fold). Genes implicated in the carnitine system for energy production that depends on fatty acid transfer were expressed significantly low (*CPT1A* = 2.25-$\log_2$ fold; *CPTII* = 2.12-$\log_2$ fold). *ACSL1* gene that was involved in fatty acid conversion was relatively stable across case-controls. Therefore, these data suggest the energy production in CRC is independent of carnitine system. Interestingly, *FGF19* expression involved in cell signaling and metabolism was significantly elevated in cases by 1.06-$\log_2$ fold which suggest the importance of this gene in energy production. Collectively, the gut microbiota-host interaction showed the increased *FGF19* expression was not associated with *NR1H4/FXR* expression, instead *FGF19* expression acts independently to regulate fatty acid oxidation for energy production. In conclusion, this study revealed that *Campylobacter* and functional content (primary bile acid biosynthesis) of the gut microbiota plays a crucial role in CRC via the independent elevated expression of *FGF19* in regulating fatty acid oxidation, not via the carnitine system.

ABSTRAK

KAJIAN TERHADAP MIKROBIUM USUS DAN PROFIL METABOLOMIK DALAM PESAKIT KANSER KOLOREKTAL

Saluran gastrousus manusia mempunyai berjuta-juta mikrobiota usus. Mikrobiota usus telah dikaitkan dengan pelbagai jenis kanser, termasuk kanser gastrik, hati dan kolorektal (CRC). Kadar CRC semakin meningkat setiap tahun dan trend ini beralih kepada umur yang lebih muda setiap tahun. Ini menandakan keperluan untuk memahami peranan mikrobiota usus dan metabolit perumah di kanser usus. Oleh itu, kajian ini bertujuan untuk menyiasat perhubungan mikrobiota usus dan menganotasi metabolit perumah yang berpotensi memainkan peranan dalam kanser usus dengan sokongan daripada kajian ekspresi gen. Sejumlah 35 sampel tisu mukosa usus kes-kawalan telah dikumpulkan dan tertakluk kepada pelbagai analisis.Tiga puluh tisu kawalan-kes telah menjalani analisis usus mikrobiota dengan menggunakan 16S rRNA penjujukan generasi seterus. Dua puluh tiga tisu kawalan kes digunakan dalam analisis metabolomik dengan menggunakan jisim spektrometri. Dua puluh lapan tisu kawalan-kes digunakan dalam ekspresi gen dengan menggunakan PCR Masa Nyata. Komposisi mikrobiota usus, metrik kepelbagaian alfa (chao1, liputan barangan, OTUs diperhatikan, indeks Shannon, indeks Simpson's, dan jarak filogenetik) dan metrik kepelbagaian beta (ANOSIM, PERMANOVA, PERMDISP dan MRPP) menunjukkan bahawa mikrobiota usus adalah stabil di pesakit kanser usus walaupun terdapat perbezaan dari segi patologi tisu (kes-kawalan). Analisis tahap genus mengenalpasti Campylobacter dari filum Proteobakteria adalah signifikan (nilai-p <0.05) lebih banyak dalam kes. Profil kandungan fungsi mikrobiota usus menunjukkan bahawa glycosyltransferases, protein biosintesis lipopolysaccharide dan kandungannya dengan ketara (nilai-p <0.05) adalah lebih tinggi dalam kes-kes.Selain itu, biosintesis asid hempedu utama adalah lebih rendah (nilai-p <0.05) dalam kes-kes yang diikuti oleh toksin bakteria serta pentosa dan interconversions glucuronate. Analisis lanjut menunjukkan Actinobacteria dan Proteobacteria merupakan phyla yang biasa menyumbang kepada enzim yang terlibat dalam biosintesis asid hempedu primer. Tambahan pula, analisis metabolomik menganotasi sebanyak sembilan metabolit polar yang diketahui dan dua yang tidak diketahui serta 19 glycerophospholipid dan sembilan metabolit bukan polar yang tidak diketahui dengan skor VIP ≥ 1.5 dan nilai-p <0.05 yang banyak dalam kes. Kesimpulan mengenai metabolit mencadangkan sistem karnitin yang memainkan peranan dalam beta-pengoksidaan untuk pengeluaran tenaga terlibat dalam CRC yang memerlukan penyiasatan lanjut dengan menggunakan ekspresi gen. Ekspresi gen NR1H4/ FXR yang terlibat dalam isyarat hempedu asid mengurang dengan ketara (3.32-log2 kali ganda). Gen yang terlibat dalam sistem karnitin untuk pengeluaran tenaga yang bergantung kepada pemindahan asid lemak telah ketara diekspres dengan rendah (CPT1A = 2.25-log2 kali ganda; CPTII = 2.12-log2 kali ganda). Gen ACSL1 yang terlibat dalam penukaran asid lemak adalah agak stabil dalam kes-kawalan. Data ini mencadangkan pengeluaran tenaga di kanser usus adalah bebas daripada sistem karnitin. Menariknya, ekspresi FGF19 yang terlibat dalam isyarat sel dan metabolisma telah meningkat dengan ketara dalam kes-kes oleh 1.06-log2 kali ganda yang mencadangkan kepentingan gen ini dalam pengeluaran tenaga. Oleh itu, interaksi mikrobium usus-perumah menunjukkan ekspresi FGF19 yang meningkat tidak berkaitan dengan ekspresi NR1H4/ FXR, malahan FGF19 bertindak secara bebas untuk mengawal pengoksidaan asid lemak untuk pengeluaran tenaga.

Oleh itu, kajian ini mendedahkan bahawa Campylobacter dan kandungan fungsi mikrobiota usus (biosintesis asid hempedu utama) memainkan peranan penting dalam kanser usus melalui ekspresi FGF19 yang mengawal selia pengoksidaan lipid, bukannya melalui sistem karnitin.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, UNITS, AND ABBREVIATIONS

ACSL1	-	Acyl-CoA Synthetase Long-Chain Family Member 1
AHR	-	Aryl Hydrocarbon Receptor
ANOSIM	-	Analysis of Similarities
APC	-	Adenomatous Polyposis Coli
bp	-	Basepair
CA	-	Cholic Acid
CAMERA	-	Collection of Algorithms for Metabolite Profile Annotation
cagA	-	Cytotoxin Associated Gene
cdt	-	Cytolethal Distending Toxin
CPT1A	-	Carnitine Palmitoyltransferase 1A
CPTII	-	Carnitine Palmitoyltransferase 2
CRC	-	Colorectal Cancer
C_T	-	Cycle Threshold
DG	-	Diacylglycerides
EMT	-	Epithelial-Mesenchymal Transition
ESI	-	Electronspray Ionization
FAP	-	Familial Adenomatous Polyposis
FXR	-	Farnesoid X Receptor
FBOT	-	Fecal Occult Blood Test
FcGBP	-	Fc Fragment of IgG Binding Protein
FGF19	-	Fibroblast Growth Factor 19
FIT	-	Fecal Immunochemical Test
FLASH	-	Fast Length Adjustment of Short Reads
GAPDH	-	Glyceraldehyde-3-Phosphate Dehydrogenase
GC-MS	-	Gas Chromatography-Mass Spectrometry
gDNA	-	Genomic DNA
GI	-	Gastrointestinal
GQN	-	Genomic Quality Number
HMDB	-	Human Metabolomics DataBase
HNPCC	-	Hereditary Non-polyposis CRC Syndromes
IBD	-	Inflammatory Bowel Disease
KEGG	-	Kyoto Encyclopedia of Genes and Genomes

KRAS	-	Kristen Rat Sarcoma Viral Oncogene
LC-MS	-	Liquid Chromatography-Mass Spectrometry
LPS	-	Lipopolysaccharide
MMCD	-	Madison Metebolomics Consortium Database
MRPP	-	Multi Response Permutation Procedure
NMR	-	Nuclear Magnetic Resonance
Nos2	-	Nitric Oxide Synthase
NR1H4	-	Nuclear Receptor Subfamily 1 Group H Member 4
OBI-Warp	-	Ordered Bijective Interpolated Warping
OTUs	-	Operational Taxonomic Units
PC	-	Phosphatidylcholine
PCoA	-	Principal Coordinates Analysis
PCR	-	Polymerase Chain Reaction
PE	-	Phosphatidylethanolamine
PE-Cer	-	Phosphatidylethanolamine-Ceramide
PEB1	-	Periplasmic Amino Acid-Binding Protein
PERMANOVA	-	Permutational Multivariate Analysis Of Variance
PERMDISP	-	Permutational Multivariate Dispersions
PG	-	Phosphatidylglycerol
PICRUSt	-	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
Q-score	-	Phred Score
QIIME	-	Quantitative Insights into Microbial Ecology
qPCR	-	Quantitative Real-Time PCR
R^2	-	Correlation
Raw PE	-	Raw 250bp Paired-End Reads
s	-	Seconds
SCFAs	-	Short Chain Fatty Acids
S.D.	-	Standard Deviation
SE	-	Standard Error
STAMP	-	Statistical Analysis of Taxonomic and Functional Profiles
TGF-β	-	Transforming Growth Factor Beta
TLR4	-	Toll-Like Receptor 4
VacA	-	Vacuolating Cytotoxin A

VIP	-	Variable Importance of Projection
XCMS	-	Various Forms (X) of Chromatography Mass Spectrometry
Δ	-	Delta
>	-	Greater Than
≥	-	Greater Than or Equal
<	-	Lesser Than
=	-	Equal
%	-	Percentage
µL	-	Microliter
µM	-	Micro Mole
ng	-	Nanogram
mg	-	Milligram
µg	-	Micrograms
rpm	-	Rotation Per Minute

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Study Background

Cancer has the mortality of 9.6 million deaths in the year 2018 (WHO, 2018). The highest cancer mortality was lung cancers, followed by colorectal cancers (CRC), stomach cancer, liver cancer, and breast cancer. In Malaysia, CRC incidence has accounted for approximately 16.4% in male and 10.7% in female (Ministry of Health Malaysia, 2019).

CRC progresses from a normal intestinal cell to cancerous cell through multiple acquisitions of genetic mutations (Vogelstein *et al.*, 1988). Genetic mutations on several genes including adenomatous polyposis coli (*APC*), Kirsten Rat Sarcoma virus (*KRAS*), Transforming growth factor beta (*TGF-β*), and finally *p53* are associated with CRC development. Despite the vast screening, treatment options, and understanding on the development of CRC, the incidence of CRC is still increasing and the incidence trend is shifting to younger age every year (Veettil *et al.*, 2017).

The human gut microbiota also known as the gut microbiome has been studied and correlated associated with various diseases including obesity, diabetes, asthma, autism, gastric cancer and also CRC (Pulikkan *et al.*, 2019; Wang *et al.*, 2017). Roughly 70% of the normal gut microbiota in human composed of *Firmicutes* and *Bacteroidetes* whereas *Proteobacteria*, *Actinobacteria* and *Fusobacteria* were observed at lower abundance (Eckburg *et al.*, 2005; Frank *et al.*, 2007). The normal gut microbiota maintains a homeostasis balance to promote a healthy feedback. Based on a previous study, dysbiosis of the gut microbiota with colonization of bacterial biofilms composed mainly *Escherichia coli* and *Bacteroides*

fragilis increases DNA damage in the colon mucosal layer that rapid onset of colorectal tumour (Dejea *et al.*, 2018). With 16S rRNA sequencing technology, the gut microbiota composition revealed greater numbers of several phyla such as *Firmicutes*, *Bacteroidetes* and *Proteobacteria* were associated with cancerous tissue in patients with CRC (Chen *et al.*, 2012). However, there is a deficit in associative studies where the biological activities between the gut microbiota and host were not addressed.

The gut microbiota and host metabolite are interrelated at the molecular and system levels that are highly complex. Metabolites such as short chain fatty acids (SCFAs) of butyrate, rather than glucose were the primary energy source for intestinal cells (Roediger, 1982). Butyrate has been reported to inhibit CRC development and promote healthy intestine by inhibiting cell proliferation, promote apoptosis, reducing tumor invasion and maintain microbial homeostasis (Wu *et al.*, 2018). Antibiotics consumption in rats has been shown to have altered intestinal metabolites by the gut microbiota and was associated with increased risk of CRC in humans (Behr *et al.*, 2018; Dik *et al.*, 2016). Therefore, what is the alteration of intestinal metabolites and their associations with the gut microbiota? Finally, these intestinal metabolites or small molecules can be investigated using metabolomics approach which could ultimately patch together the pathways that are involved in CRC tumorgenesis which are later validated using gene expression analysis.

1.2 Problem Statement

CRC incidence has been increasing and approximately 60% of cases are detected at the later stages (Stage III and IV). Accumulating evidence has shown that the gut microbiota was associated with CRC. However, association studies have not demonstrated the gut microbiota-host interactions and biological activities in CRC.

1.3 Hypothesis

There are differences in the gut microbiota and their activities in between tumour and adjacent normal mucosal in CRC patients.

1.4 Objectives

The main aims of this study are:

a) To determine the intestinal microbial alpha and beta diversities in CRC patients.

b) To identify significantly different intestinal microorganisms and determine the functional pathway in between CRC and normal colon mucosal.

c) To annotate significantly different metabolites from CRC and normal colon mucosal tissue.

d) To propose pathways that might contribute to in CRC based on functional pathway analysis and annotated metabolites.

e) To validate the involvement of the pathway in CRC using gene expression analysis.

1.5 Scope of Study

This study focuses on investigating the gut microbiota-host interactions and biological pathways in CRC as mentioned in the objectives of the study. Other etiological variables would not be covered in the present study.

This limitation is due to the fact that it would be too great of a scope to cover the etiology in CRC. Numerous findings on various CRC related etiologies such as antibiotic consumption, dietary consumptions, family cancer history and smoking habits were shown to be associated with CRC (Carr *et al.*, 2018). Although it is interesting to investigate the biological pathways of several different types of etiology with CRC, but it is deemed to be too wide to be covered in this study. Apart from that, it is even more crucial to establish the underlying basis of gut microbiota-host interaction and biological pathways involved in CRC before embarking into the studying associations of various etiologies with CRC.

Therefore, this study aims to investigate the gut microbiota-host interactions and the biological pathways in CRC by studying the gut microbiota, metabolomics, and gene expression in CRC and adjacent normal mucosal tissue. The gut microbiota and its diversity were determined using 16S rRNA sequencing followed by functional content profiling of the gut microbiota. Following that, significantly elevated metabolites in CRC and adjacent normal mucosal tissue were determined

and annotated using mass spectrometry. From the metabolomic data, potential pathway involved in CRC was proposed and further confirmed using gene expression analysis.

CHAPTER 2

LITERATURE REVIEW

2.1 Cancer Statistics

Cancer is a major public health concern and the leading cause of death in the world. Cancer mortality has been increasing since the year 2010 at 8.1 million deaths to 9.6 million deaths in the year 2018. The highest cancer mortality was lung cancers (1.8 million deaths) followed by colorectal cancer (CRC) (0.88 million deaths), stomach cancer (0.78 million deaths), liver cancer (0.78 million deaths), and breast cancer (0.63 million deaths) (WHO, 2018).

In Malaysia, cancer has accounted for ≈ 24,000 of deaths in the year 2015. respiratory cancer have the highest mortality (≈ 4600 deaths) followed by breast cancer (≈ 2800 deaths), CRC (≈ 2600 deaths), liver cancer (≈ 1900 deaths), and lymphomas, multiple myeloma (≈ 1200 deaths) (WHO, 2016). Recently in 2019, there were 16.4% and 10.7% of CRC incidence for Malaysian male and females, respectively (Ministry of Health Malaysia, 2019).

2.1.1 Incidence and Mortality in Colorectal cancer (CRC)

CRC is the third leading cause of death worldwide as well as in Malaysia with 0.77 million deaths and ≈ 2600 deaths, respectively. In the year 2012, there was an estimated 1.4 million newly diagnosed CRC cases worldwide (Ferlay *et al.*, 2012). There are three forms of CRC namely sporadic which accounted for more than 85% of all CRC cases, familial form that constitute approximately 10% of all CRC cases which include familial adenomatous polyposis (FAP), and hereditary non-polyposis CRC syndromes (HNPCC). The highest CRC incidence was in China (0.25 million cases), followed by USA (0.13 million cases), and Japan (0.11 million cases). Malaysia has an estimated 4539 CRC cases in the year 2012 and ≈ 70% of the

cases was diagnosed at the advance stages (stage three and stage four) (Shah *et al.*, 2014).

The survival rate of CRC patients are closely correlated with the stage of cancer at the time of diagnosis. The 5-year survival rate for early cancer stages are 90% for stage 1 and 80% for stage 2. Unfortunately, the 5-year survival rate for cancer diagnosed at advance stages are only 30-60% for stage 3 and 5-10% at stage 4 (Haggar *et al.*, 2009; Sankaranarayanan *et al.*, 2014). It is important to note that the need to perform emergency surgery on patients with advanced stage CRC presented with obstructive symptoms was a major cause of poor CRC survival in advance stage CRC (Stillwell *et al.*, 2011).

Current medical technologies have paved new and innovative ways to detect early stages of CRC using non-invasive and invasive approaches. Presence of blood in fecal is one of the classical symptoms of CRC which can be detected through direct visualization or fecal testing (Mansson *et al.*, 1999). A non-invasive blood fecal detection method used in CRC screening known as fecal occult blood test (FOBT) has been reported to reduce CRC related mortality (Levin *et al.*, 2008; Sung *et al.*, 2015), but a dietary restriction is required before test, whereas fecal immunochemical test (FIT) does not require dietary restriction and offers higher specificity is always preferred (Ouyang *et al.*, 2005; Redwood *et al.*, 2014). Besides, a validated seven-gene biomarker panel (*ANXA3*, *CLEC4D*, *TNFAIP6*, *LMNB1*, *PRRG4*, *VNN1*, and *IL2RB*) analyzing the gene expression level of blood in CRC patients as compared with controls had a 77% specificity, 61% sensitivity, and 70% accuracy (Yip *et al.*, 2010). Upon suspicion of potential CRC in subjects using non-invasive screening approach, colonoscopy is needed to obtain a biopsy specimen for diagnosis confirmation followed by imaging for CRC staging.

Colonoscopy is still the main procedure for CRC diagnosis given to its precision and effectiveness (Lauby-Secretan *et al.*, 2018; Levin *et al.*, 2008). However, the procedure is invasive and requires skilled healthcare personnel and sufficient resources. Colonoscopy might not be applicable to be used as a tool for CRC screening in a large population of Malaysia, where there is insufficient specialist to attend the large number of subjects undergoing the procedure (Hwong

et al., 2014). Hence, prioritizing high risk subjects with positive results from non-invasive CRC methods to undergo colonoscopy could allow efficient allocations of healthcare resources.

Despite the multiple screening and treatment strategies put in place to tackle CRC, the incidence and mortality of CRC still occurs at an alarming rate and often diagnosed at an advance stage. Hence, it is imperative to understand the associated risk factors, carcinogenesis and progression of CRC in the Malaysian population to tailor a suitable prevention and treatment programs.

2.2 Carcinogenesis and Progression of CRC

CRC develops through a sequential order of events which begins with the transition of normal colon epithelium to adenomatous intermediate and then finally cancer also known as adenocarcinoma (Figure 2.1). Multiple genetic abnormalities especially genomic instability is required for tumour progression. This multi-step CRC carcinogenesis process has been introduced by Vogelstein *et al.* (1988). The study proposed that mutations in a tumour suppressor gene known as adenomatous polyposis coli (*APC*) gene is the initial step of the carcinogenesis process followed by Kirsten Rat Sarcoma virus (*KRAS*), transforming growth factor beta (*TGF-β*), and *p53* mutations. A separate study has reported that majority of adenomatous polyps have of *APC* mutations in both alleles which constitutes approximately 85% of sporadic CRC tumours and also involved in familial adenomatous polyposis (FAP) (Smith *et al.*, 2002).

APC gene has been involved in the regulation of β-catenin, apoptosis, cell cycle regulation, cytoskeleton organization and cell adhesion (Sieber *et al.*, 2000). Translated APC protein from *APC* gene is a main player for Wnt signalling pathway that regulates cell proliferation through β-catenin (Teo & Kahn, 2010). The binding of normal APC protein to β-catenin leads to the degradation of β-catenin which halts cell proliferation signals in Wnt signalling pathway (Najdi *et al.*, 2011). However, mutant APC protein cannot bind and degrade β-catenin that leads to the translocation of β-catenin into the nucleus and binds to a transcription factor which activates cell proliferation and carcinogenesis process (Tetsu and McCormick, 1999).