1.0: Introduction

Cancer is an abnormal biological condition that is characterized by unchecked growth of cells in virtually any parts of the body. These cells would accumulate and form a cell mass that is detrimental to the body because they not only can invade and destroy the surrounding tissues as they expand, the cells can also migrate from the initial site of growth and spread to other body parts. The destructive activity of these cells results in failure of normal physiological processes in the body and eventually leads to death. This disorder is a major health problem plaguing all populations worldwide in the recent decade. Cancer was the leading cause of disease-related death in 2004 (13% of total disease-related death). In the latest epidemiological study however, cancer is second to cardiovascular diseases in causing disease-related death around the world (World Health Statistics, 2008). The ever-increasing number of cancer-related death cases reported per year is alarming and the estimation is on course to meet the forecast of 11.4 million cancer-related death worldwide by 2015 and 12 million by 2030. Among all cancer types, lung cancer is the leading death-causing cancer in the world followed by stomach, colorectal, liver and breast cancers (World Health Statistics, 2008).

Genetic factor has been implicated in the initiation of cancer. Carcinogenesis is attributed to gene mutations that are acquired either sporadically or by inheritance. Inherited cancer-causing mutations, in particular, would affect the body by causing intermediate diseases which, over time and in the absence of any clinical intervention, would lead to cancer. Among the diseases that have been linked to increased susceptibility to cancer are Ataxia telangectasia, Bloom syndrome and Fanconi's anemia (leukemia), Li-Fraumeni syndrome (cancer of the bone and soft tissues), Cowden syndrome (cancer of the thyroid and breast) and Peutz-Jeghers syndrome (hamatomas). In this mutation screening study, the focus was set on Familial

Adenomatous Polyposis or FAP, which is a syndromic form of colorectal cancer. Colorectal cancer develops when epithelium cells lining the intestines began to divide excessively. In this context, a person with FAP will start to develop polypoid growth along the intestinal lumen that could turn cancerous over time, predisposing the patient to colorectal cancer. Often, early detection of these predisposing conditions by regular medical check-up and genetic testing can prevent disease progression to cancer. As such, the outcome of this study is hoped to benefit genetic testing efforts by enriching information on mutation spectrum of Malaysian FAP patients.

As discussed in an article by Mohandas (2001), studies in genetic predisposition to cancer can also help in the understanding of carcinogenesis. The effect of gene-environment interaction in carcinogenesis can be assessed, in particular, by using patients with inherited cancers that attend cancer family clinics as human model in epidemiological study sets. Furthermore, fundamental findings from research in genetic predisposition to cancer would also support advances in molecular cancer research. Heritable cancer can serve as a model to explain the carcinogenesis of its corresponding sporadic type as they would have similarly affected genetic pathways. Cancer predisposition knowledge could also be applied to shed more light on cell growth regulation studies as genetic alteration predisposing to cancer may be linked to failed cell cycle control mechanism (Mohandas, 2001).

In Malaysia, information on FAP incidence is scarce despite the existence of colorectal cancer management and database. Progress of scientific research on FAP-colorectal cancer in Malaysia is rather lagging, as evident in the low number of publications (clinical and/or molecular reports) on the disease that are available online to date. This reflects the huge gap that exists in the knowledge of genetic predisposition to colorectal cancer in Malaysia. It has become obvious how knowledge in genetic predisposition to cancer can benefit cancer prevention efforts and treatment plans in an

early-onset type disease like FAP. This calls for more investigations and researches to be conducted in this area to facilitate better understanding of the disease and its prevention.

2.0: Literature review

2.1: Cancer

2.1.1: Cancer definition

Normal cells undergo cell division to serve numerous physiological purposes such as cell specialization, growth, reproduction, immune defence as well as cellular repair. It is a multistep process that involves a cell entering tightly regulated phases of growth to produce daughter cells of the exact same copy of genetic information. These daughter cells will repeat the same cycles of division to give rise to more cells.

However, cells do not have the ability to perpetuate and are often compelled into cell death. It is essential for cell proliferation process to be balanced by programmed cell death in order to maintain cell number equilibrium at any given time. But cells can gain the ability to override the regulatory factors controlling its replication and homeostasis processes. The ability to take on unrestricted cell reproduction together with evasiveness towards cell death directives makes these cells potentially immortal, and thus affecting the balance between cell proliferation and cell death.

Cancer is a term widely used to describe cells that are in this state of unchecked division and perpetuation, although not all abnormally dividing cells are cancerous. Excessively proliferating cells that form tumours or neoplasm are essentially self-limited and non-detrimental. These harmless neoplasms are considered as benign tumours and pose no threat to life unless the cells that form the tumour have specific function in physiological processes like producing hormones or the cell mass is physically obstructing and compressing vital organs. Cells that form benign neoplasm maintain their morphology similar to the type of tissue they were derived from and do not exhibit local tissue infiltration and me tastatic capabilities. They grow in an expansive manner, forming circumscribed and often encapsulated mass, but may eventually regress over time (Mellor, 2006).

Cells from malignant neoplasm, on the other hand, are cancerous. Cancer is characterised by the ability of cells to not only replicate unceasingly but also to migrate and populate other parts of the body. Moderate and poorly differentiated malignancies have cell morphology that is increasingly dissimilar to their tissue of origin, especially in anaplastic tumours. Cancerous cells grow rapidly in numbers and spread locally by invading their neighboring tissue and systemically through distant metastasis (Tanaka, 2009).

2.1.2: Carcinogenesis

According to a review by Hanahan and Weinberg (2000), for a cell to become malignant it has to overcome the following six major physiological hurdles; acquiring limitless replication potential, self-sufficiency in nutrients and growth factors, ability to remain impervious to anti-growth factors, resisting programmed cell death directives, ability to stimulate and sustain angiogenesis as well as to gain invasive and metastatic capabilities. Most cancers, if not all, require all six hallmark capabilities to survive. While it is believed that a single anomalous regulatory system would affect other interrelated pathways, no specific order or successive pattern was observed for tumour cells to acquire carcinogenic traits (Hanahan and Weinberg, 2000).

Cancer has been hypothetically linked to genetics for as long as a century and a half ago. A publication by Hansen and Cavenee in 1987 reviewed that inheritance of malignancy was observed as early as in 1866. And since then, a myriad of scientific reports have been produced with evidence that support this relationship. Carcinogenesis occurs as a result of accumulated mutations that are acquired either by sporadic or inheritance in the oncogenes and tumour suppressor gene and producing dominant gain of gene function or loss of gene function respectively (Bishop and Weinberg, 1996).

Tumorigenesis begins with the clonal proliferation step from a single mutated cell. Fearon *et al.* (1987) investigated the clonal composition of tumours from the early stages of growth and reported that most tumours have a monoclonal composition reflecting the cell it arose from. Clonal expansion is thought to help in creating a sufficiently large target for one of the cells to accumulate the next needed mutation, especially in diseases that require multiple mutations (Brash, 1997).

Tanaka (2009) in his review on human colorectal carcinogenesis highlighted that there are two proposed models to explain the development of colonic adenoma based on histochemical observations. The first is the 'top-down' model which states that mutant cells emerge from one of the cells that make up the surface of the lumen. This is supported by histochemical evidences that demonstrate the progression of the adenoma cells from the intercryptal zone to the basal crypt through clonal expansion (Shih *et al.*, 2001) (Figure 2.1). Conversely, clonal expansion of adenoma cells has been shown to follow a 'bottom-up' model in which a defective stem cell with proliferation-activating mutations from the basal depth of the crypt moves into the intracryptal zone and expands from this position (Preston *et al.*, 2003). The mechanistic exclusivity of both carcinogenesis models however is still controversial and ultimately adds another layer to the complexity of colorectal neoplasia initiation mechanism.

2.1.3: Types of cancer

The organs in the human body consist of different types of cells and virtually all types of cells can develop into cancer. As such, cancer is classified according to the type of cell it derived from rather than its site of origin. Cancer that developed from mesenchymal cells is called sarcoma while cancer of the epithelial cells is termed as carcinoma. On the other hand, hematopoietic cells that have turned malignant are

commonly referred to as lymphoma while a tumour that resembles embryonic tissues is called a blastoma.

Similar to cancers of the lung and breast, cancer that occurs in the colon and rectum is characterised as carcinoma because it develops from the epithelial cells that line the mucosal layer of the intestine. This disease affects the large intestinal tract and the rectum by the formation of adenomatous polyps and hyper-inflammatory lesion of the mucosal lining.

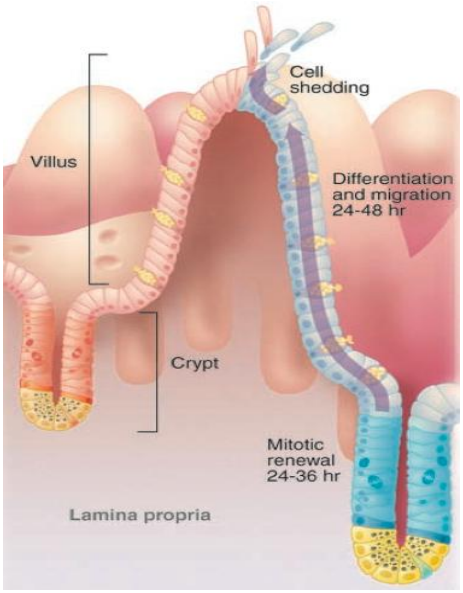

Figure 2.1: The cross-section depiction of the intestine's epithelial layer. The protruding villus is lined by a layer of epithelial cells that absorbs amino acids, sugar and vitamins from the digested food. Loss of cells by epithelial shedding around the villus tip is counter-balanced by the production of new epithelial by the stem cells (blue) located next to the paneth cells (yellow) close to base of the crypt. Progenitor stem cells will undergo differentiation to form absorptive epithelial cells, goblet and entero-endochrine cells. They are pushed up towards the villus tip to compensate the cell loss by epithelial shedding.

(Source: Radtke and Clevers, 2005)

2.1.4: Anatomy of colon and rectum

The human intestinal tract is composed of duodenum, the small intestine (jejunum, ileum and cecum) and the large intestine called colon. The colon can stretch up to 150cm in length and has four parts; the ascending, transverse, descending and sigmoid colon (Figure 2.2). The rectum is the end part of the colon that is connected to the anal valve. It is approximately 12cm long and functions as fecal storage compartment prior to defecation.

The human colon is a hollow structure that is lined by a mucousal layer. This layer is made up by epithelial cells that absorb water and electrolytes from the passing feces. A layer of muscles comprised of circular and longitudinal muscles wraps the mucousal layer to help move the feces along the intestine. The outer surface of the colon is covered with a peritoneal layer. Apart from absorbing solutes, the colon also houses digestive bacteria that help to break down residual undigested food and produce fat-soluble vitamins that can be further absorbed by the host.

(a) **(b)**

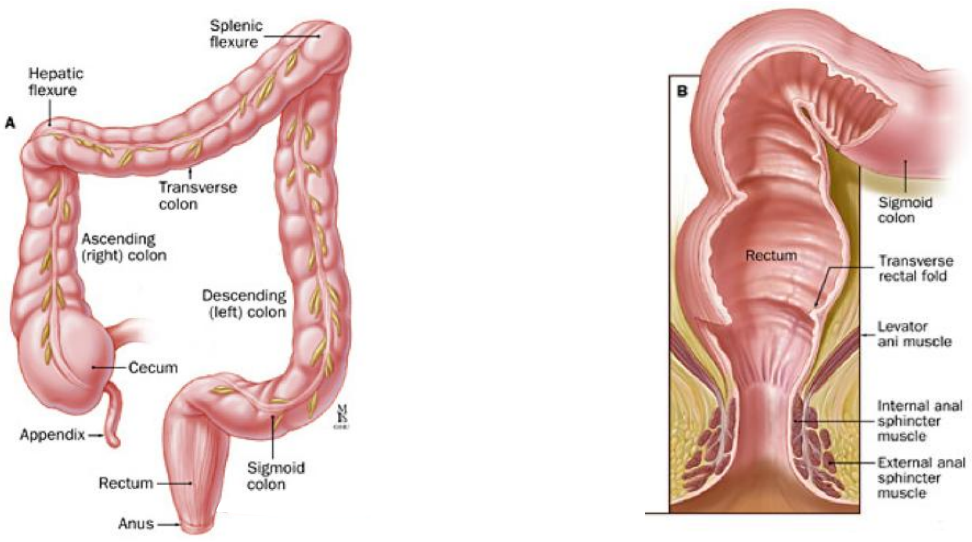

Figure 2.2: Normal anatomy of the human (a) colon and (b) rectum.
(Source: www.hopkins-gi.org)

2.2: Colorectal cancer

2.2.1: Epidemiology

As the name implies, colorectal cancer (CRC) is the type of cancer that affects the large bowel. This cancer can be found at any spot along the ascending, transverse and descending colon as well as the rectum. It is more common in populations of developed countries largely due to Westernized lifestyle and dietary habits compared to developing and third world countries (Marchand *et al.*, 1997; Slattery *et al.*, 1999).

In general, colorectal cancer recorded the fourth highest number of cases for cancer-related death in the world and it is the third highest cancer diagnosed in 2002. With 655,000 deaths worldwide per year, it is the third leading cause of cancer-related death in the western countries (Cancer, 2010; World Health Statistics, 2010).

In Malaysia, it is the third highest type of cancer diagnosed in males and females, and when combined they make the highest number of cancer cases diagnosed, reaching 3600 cases yearly. In addition, the incidence data showed inclination towards ethnical background with the Chinese population in Malaysia showing higher incidence of colorectal malignancy compared to other main racial groups, namely the Indians and Malays (Lim & Yahaya, 2004).

The incidence of CRC among Asian populations, however, is not different when compared with Western countries in terms of its anatomical distribution and mortality. Colorectal cancer cases in many Asian countries have risen two- to four-folds in the past decades, with colorectal non-polypoidal neoplasmic form being the most common type. This complicates the early screening efforts for cancer prevention in such patients because polypoidal neoplasms are much easier to detect (Sung *et al.*, 2005).

2.2.2: Aetiology and risk

There are several factors that increase the risk of developing colorectal cancer, which include age, race and ethnic background, family history of colorectal cancer as well as prolonged inflammatory bowel syndromes. Daily lifestyles also contribute to increased cancer risks such as smoking and alcohol consumption, diets with high fat content, red meat and processed meat intakes, lack of exercise as well as being overweight and diabetic (Cancer, 2010).

Colorectal cancer occurs sporadically or by genetic inheritance. About 20-25% of total colorectal cancer cases show familial pattern. However, it is difficult to ascertain if a case is truly familial or has occurred by chance, except in cases with clear familial colorectal syndromes. Likewise, sporadic cases may be falsely diagnosed due to unavailability of family medical background, small family size, and poor diagnostics (de la Chapelle, 2004).

The colorectal cancer risk of an average person above 50 years of age with no familial colorectal cancer history is low (5-6%). The risk is however increased four-folds if first- and/or second-degree family members were diagnosed with colorectal cancer (Rustgi, 2007).

2.2.3: Diagnosis and treatments

Colorectal cancer diagnosis involves screening techniques in the form of stool testing or direct bowel examination. Fecal matter can carry a telltale sign of neoplasmic growth in the form of blood stains. This test can provide simple and inexpensive screening for occult blood in stool samples with up to 90% sensitivity level.

Bowel examination can be carried out either using endoscopic diagnosis or barium edema. Endoscopic examination uses flexible fiber optic tube incorporated with a camera to visualize the lumen of the colon. Two types of endoscopy, namely

sigmoidoscopy and colonoscopy are used for examination with the latter being more favourable in most screenings due to its high diagnostic sensitivity and specificity. Barium enema, on the other hand, requires the use of barium as the contrasting agent to visualize the insides of the colon. The colon will be filled with barium solution before being observed under X-ray through a monitor. Detection through double contrast radiology has high sensitivity range (39-90%) but entirely dependent on the ability of the radiologist to identify even the most subtle lesion (Johns Hopkins Gastroenterology and Hepatology, 2010).

In brief, treatment of colorectal cancer involves therapies with different degree of invasiveness. Small (less than 5mm diameter) polyps can be removed by polypectomy during colonoscopy. Advanced tumour growth, however, requires surgical therapy through colon resection. Depending on the location of the tumour, removal of the ascending colon (right hemicolectomy), transverse colon (transverse colectomy), descending colon (left hemicolectomy) or the sigmoid colectomy can be carried out. Tumour located in the rectum, on the other hand, requires treatment actions that range from local excision to abdominoperineal resection, again depending on the exact position of the growth (Johns Hopkins Gastroenterology and Hepatology, 2010).

2.3: Familial Adenomatous Polyposis (FAP)

Familial Adenomatous Polyposis or FAP is a syndrome that predisposes to colorectal cancer. This syndrome, also known as Familial Adenomatous Coli, is characterized by the formation of polypoidal growth throughout the colon and rectum. This condition is first described by Corvisart in 1847 and was later given a more definite account by Chargelaigue in 1859 (Spigelman, 2004).

FAP affects approximately 1 in 10000 people in the European Union (Public summary of positive opinion for orphan designation of celecoxib for the treatment of

Familial Adenomatous Polyposis [FAP]). As such, FAP is considered as a rare disease. It is an early onset disease with an average age for polyp development as early as 16 years old, and malignancies that can arise at the mean age of 40 (Neklason *et al*., 2004). In classical form of FAP, polyp count can range from hundreds to thousands. FAP however has a milder counterpart known as attenuated FAP or AFAP. Compared to FAP, AFAP exhibits lower polyp count (normally less than one hundred polyps). Nevertheless, one or more of these polyps can progress and transform into malignant neoplasm if no early detection screening and preventive surgical intervention were administered.

Some of the common symptoms identifiable to FAP include abdominal pain, rectal bleeding and diarrhoea. At the same time, FAP patients are likely to develop extracolonical manifestations such as benign lesions of the retina and bones as well as development of desmoid tumours and polyps in the upper gastrointestinal tract (Gebert *et al*., 1999; Lal and Gallinger, 2000).

2.4: *Adenomatous Polyposis Coli (APC)* gene

2.4.1: *APC* gene discovery

To find candidate genes for cancer, the search should begin by looking at genes with *a priori* basis for expecting an effect such as mutations that adversely affect protein functions (Bodmer, 2006). In colorectal cancer, many molecular techniques were employed to identify the potentially corrupted causal gene. For example, *ras* gene was found to be oncogenic in colorectal cancer through transfection assay that demonstrated how mutated *ras* allele confer recipient cells with neoplastic properties (Barbacid, 1987). RFLP is employed to detect loss of gene heterozygosity (LOH) of polymorphic DNA markers by comparing restriction pattern of patient's normal and tumour tissues (Vogelstein *et al.,* 1989).

Cytogenetic examination of chromosomes in FAP patients identified an anomaly in chromosome 5, in particular, a constitutional deletion of chromosomal band 5q21 (Herrera *et al.,* 1986). By cloning a large portion of 5q21 into yeast artificial chromosome (YAC) followed by chromosome walking, Joslyn and colleagues (1991) found a 100-260kb deletion believed to be part of the *APC* gene. To identify what gene was lost due to the deletion, various cDNAs was tested to probe 5q21 chromosomal region, through which cDNA from the *APC* gene was positive for the detection. Further analysis of the *APC* gene by PCR and DNA sequencing revealed various germline mutations such as nonsense and missense mutations (Nishisho *et al.,* 1991) as well as point mutations that cause frameshift (Groden *et al.,* 1991). *APC* gene was also found to be the target for somatic mutation in sporadic colorectal cancer cases (Nishisho *et al.,* 1991).

An *APC* gene homolog in mouse model was established through mutagenesis studies of murine intestinal neoplasia (*Min*) (Su *et al.,* 1992). The murine *APC* gene homolog has proved useful in FAP studies as it provided an animal model for studies of environmental and chemical factors towards carcinogenesis as well as assisting in development of diagnostic and therapeutic strategies. Other knockout mice strain such as Apc1638N and *Apc(716)* have since been developed by the same gene targeting method (Fodde, 1994; Oshima *et al*., 1995).

2.4.2: Background of *APC* gene

APC is a large gene consisting of 15 exons. It is located at position 5q21 of chromosome 5 and comprises 8535 nucleotide pairs that codes for a complete APC protein of 2834 amino acids (Kinzler *et al*., 1991; Groden *et al*., 1991). All the exons are generally around 100-300bp long except for exon 15 which is 6577bp. Exon 15 alone contributes to 75% of the APC protein size.

To date, there are more than 900 germline mutations found ubiquitously throughout the gene that can produce 597 variants of the APC protein (Universal Mutation Database; Beroud *et al.,* 2000). Point mutations are the most common type of mutations to cause FAP, most of the time generating stop codons and the eventual truncation of the APC protein (Beroud and Soussi, 1996). Because of its long coding region, exon 15 has become the common target for germline and somatic mutations. Exon 15 contains a region known as mutation cluster region (MCR). This region is located specifically between codon positions 1250-1464 and accounts for 60% of all somatic mutations (Nakamura *et al*., 1991). Codons 1309 and 1061 of exon 15 are found to have high frequency of mutations and these 'hotspots' account for 17% and 11% of total germline mutations reported, respectively.

The position of pathogenic mutations are not restricted to the coding regions *per se*. Alteration in the non-coding DNA regions such as deletion in the gene promoter (Charames *et al*., 2008) as well as intronic sequences that define intron-exon boundaries (Fostira *et al*., 2010) have been shown to cause *APC* gene inactivation. Similarly, silent mutations in the coding regions may contribute to *APC* gene malfunction if the mutations occurred in splicing regulatory sites. In such events, mutations will give rise to aberrant splicing products that might contribute to the variable expression of the disease (Cartegni *et al*., 2002).

FAP is a hereditary autosomal dominant disease with 100% penetrance level. However, the phenotypic spectrum of FAP is rather broad and interestingly, was discovered to have a close correlation with the positions of mutation (Nieuwenhuis and Vasen, 2007). Mutations between codons 1250 to 1464 are often implicated with profuse polyposis while mutations in the regions between codon 157-1249 and 1465-1595 were found to cause intermediate polyposis. Mutations within the 5' and 3'

regions of the gene on the other hand are linked with the attenuated version of FAP (Table 2.1).

Table 2.1: Classification of FAP severity (Nieuwenhuis and Vasen, 2007).

Type	Phenotype	Codon position of mutation in the gene	No. of colorectal adenomas	Age of onset (decade)
Classical	Profuse	1250-1464	Thousands	1st and 2nd
	Intermediate	157-1249,1465-1595	Hundreds to thousands	2nd and 3rd
Attenuated	Attenuated	5' and 3' end of gene	Less than 100	4th and 5th

2.4.3: APC protein functions

APC protein was found to be an integral component in the activation of the Wingless-type MMTV (mouse mammary tumour virus) integration site family (Wnt) signaling pathway. Wnt signaling is an evolutionary conserved pathway that controls many events during embryogenesis, regulating the expression of the genes involved in morphology, proliferation, mediation of intercellular adhesion and motility as well as cell fate processes (Sieber *et al*., 2000; Thorstensen and Lothe, 2003).

APC protein works in concert with Axin protein and glycogen synthase kinase (GSK)-3b to negatively regulate β-catenin cytosolic level during deactivated Wnt signaling state (Spink *et al.*, 2000). In the absence of Wnt pathway-activating ligand, this multiprotein complex binds and phosphorylates β-catenin for protein degradation. However, upon Wnt signaling activation, the β-catenin phosphorylation was inhibited by the activation of the protein dishevelled (Dsh/Dv1) that blocks the priming of (GSK)-3b to β-catenin. As a result, β-catenin become stabilized and accumulate in the cytosome and translocate into the nucleus, facilitating the activation of Wnt target

genes, among others, *cyclin-D1*, *Axin2/conductin* and *c-myc* (Amit *et al.*, 2002; Rustgi, 2007). A simple representation of the Wnt signaling pathway is shown in Figure 2.3.

Truncated APC protein without its protein binding domains would fail to form the β-catenin-targeting multiprotein complex, causing unnecessary accumulation of cytosolic and nuclear β-catenin. Elevated levels of β-catenin contribute to cell hyperproliferation, deeming β-catenin as a proto-oncogene and *APC* gene as a tumour suppressor gene (Spink *et al.*, 2000).

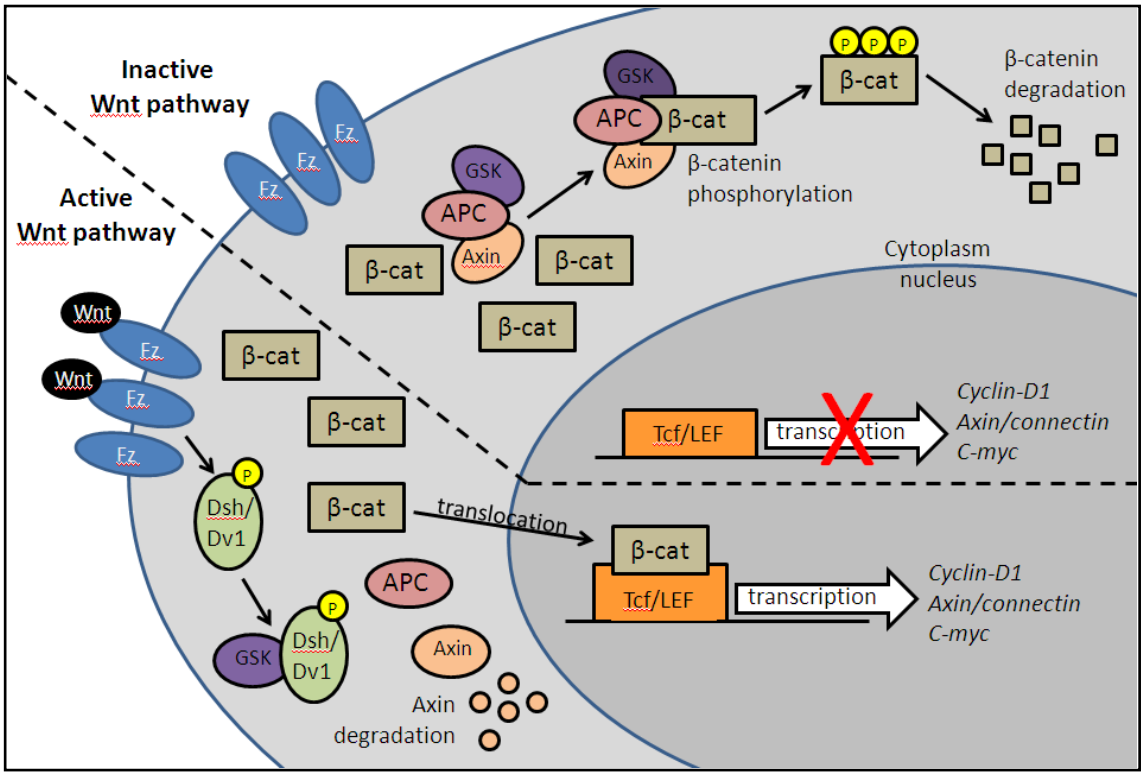

Figure 2.3: Wnt signalling pathway. This pathway controls the expression of multiple target genes by tightly regulating the cytoplasmic level of β-catenin (β-cat). Defect in the function of the APC-GSK-Axin complex thus would contribute to constant activation of the pathway (Herbst and Kolligs, 2007).

2.5: Mutation screening

2.5.1: Methods for mutation screening

Mutation screening and detection of clinical samples is a crucial process in presymptomatic genetic testing of heritable diseases. Mutation screening of patients with a familial disorder as well as high-risk immediate family members are carried out to identify the causative mutation and how it segregates within the family. At present, many quick and precise screening methods are available for use in genetic testing such as direct DNA sequencing, denaturing gradient gel electrophoresis (DGGE; Myers *et al.*, 1987), *in silico* single-strand DNA analysis known as high resolution melting analysis (HRMA; Vossen *et al.*, 2009), enzymatic and chemical cleavage methods (del Tito *et al.*, 1998; Cotton *et al.*, 1988), protein truncation test (PTT; Roest *et al.*, 1993), heteroduplex analysis (HDA; Nijbroek *et al*, 1995) and single-strand conformation polymorphism (SSCP; Hayashi, 1991, Peltonen *et al.*, 2007).

2.5.2: Single-strand Conformation Polymorphism (SSCP)

SSCP was first described by Orita and colleagues in 1989 to compensate the shortcomings of other methods for detection of sequence variation. SSCP method was developed based on the feasibility of using mobility shift due to DNA conformation change to detect polymorphisms and point mutations. The process consists of several simple steps starting with PCR amplification of DNA fragments followed by heat denaturation and snap-chill reannealing. DNA samples are then separated on a native matrix gel based on its conformational change. DNA sequence variation amongst samples would translate into band mobility shift.

2.5.2.1: Critical parameters of SSCP

Several optimisation steps were done to improve the sensitivity and simplicity of the technique by modifying the experimental condition of SSCP analysis. Among the critical parameters for an effective SSCP are electrophoresis condition (gel temperature and ionic concentration of electrophoresis buffer) and length of DNA fragments for analysis (Hayashi, 1991; Kakavas *et al*., 2008).

DNA fragments less than 400bp in size should be used in order to get the best result of SSCP (Hayashi, 1991). The ability to detect band shifts is close to 100% when fragments of around 150-200bp length are used (Prosser, 1993; Sheffield *et al*., 1993, 1998). However, percentage of shift detection sensitivity decreased when longer fragments (up to ~1300bp) are used (Hongyo *et al*., 1993; Fan *et al*., 1993; Sheffield *et al*., 1993). Analysis of large DNA fragments can be improved either by optimisation of the electrophoresis conditions (Fan *et al*., 1993) or by DNA fragment digestion to produce smaller sizes using restriction enzyme. However, this approach depends on the availability of enzyme cut site in the DNA fragment of interest (Hayashi, 1992).

The most critical parameter in SSCP analysis is the temperature of the gel during SSCP electrophoresis. Conformational folding of single-stranded DNA into secondary structures is mediated by intramolecular hydrogen bonds between nucleotides in the DNA strand (Jackel *et al.*, 1999). These bonds are temperature sensitive and depends on the thermal stability of the complementary nucleotide stretches. Since mutation screening is done by observing shifts in mobility, it is therefore important to maintain a constant temperature during electrophoresis to preserve the single-strand DNA conformations in order to improve band sharpness and reproducibility of strand separation (Hongyo *et al.*, 1993).

Ionic components in buffer solutions play a part in the formation of higher-order structure of the single-stranded DNA. Suboptimal electrical conductivity due to ionic

concentration may increase heat production (Orita *et al.*, 1989; Hayashi and Yandell, 1993), leading to undesirable fluctuation of temperature and oscillation of DNA conformation (Fan e*t al.*, 1993).

Among the methods used to maintain constant running temperature include the use of water-jacketed gel, cooling fans, carrying electrophoresis in a cold (4°C) room as well as using cooling water-circulating system (Fan e*t al.*, 1993; Hayashi and Yandell, 1993; Hongyo *et al.*,1993). Alternatively, adding 5-10% of glycerol into the gel can give as good a result as cold SSCP would show (Hayashi, 1992). This is because glycerol can lower down the pH of the gel. This will suppress the charge of the phosphate backbone of the DNA, stabilising the single-stranded conformation (Myers *et al.*, 1998).

2.6: Objectives of study

This study has two main objectives. The first objective is to screen the *APC* gene for mutations that might be pathogenic in a set of patients with Familial Adenomatous Polyposis (FAP) using PCR-SSCP and direct sequencing method. The second objective is to investigate whether the c.847C>T mutation located in a putative exonic splicing enhancer (ESE) sequence in exon 8 would result in exon skipping.

This second study is a continuation of a previous investigation that seeks to explain why a 55 year-old Malaysian Chinese heterozygote carrier of this mutation did not present any FAP symptoms and was later diagnosed as ‘unaffected’ (see Mohamed *et al.* [2003]). The author suggested that the c.847C>T mutation may abolish a splicing enhancer binding site that would cause skipping of exon 8 from the mature transcript. Because the resulting reading frame would still be retained during translation, the end protein product would be expected to be a functional isoform deficient in only 23 amino acids compared to the original protein.

Analysis carried out *in silico* predicted that the mutation would indeed abolish an ESE in exon 8 and could cause the exon to be skipped during splicing. However, since mRNA sample from this individual was not available for analysis, an *in vitro* splicing assay was designed in this study to confirm the prediction. Specific aims for this assay are as follows:

1) To introduce a c.847C>T mutation into exon 8 of the *APC* gene construct.
2) To construct an *APC* minigene that consists of exon 7, 8 (mutant and normal) and 9.
3) To clone the *APC* minigene into a mammalian expression vector.
4) To transfer and express the minigene in a mammalian cell culture system.
5) To isolate and analyse mRNA products expressed from the *APC* minigene for splicing abnormalilty.

3.0: Methodology

3.1: Study subjects and samples

The study subjects are comprised of an Familial Adenomatous Polyposis (FAP) family designated as FAP-06 and two FAP patients with no available information on family background. They are referral cases received from Selayang Hospital in Kuala Lumpur. Blood samples of the patients were kept in -20°C for long-term storage.

The pedigree of family FAP-06 is shown in Figure 3.1. A member of this Malay family (designated as 06-1) was diagnosed with FAP after he was found to have well-differentiated adenocarcinomas in his rectum as well as severe dysplastic adenomas in his periampullary tract. He was 66 years old when diagnosis was made. Familial segregation of colorectal cancer was evident in the family with the passing of the patient's mother at age ±70 and two of his brothers (at age ±50 and ±45) due to colorectal cancer. Patient 06-1 had had two sons, both of which have passed away due to unconfirmed cause of death. Both sons of 06-1 were survived by altogether six children. Blood samples from all six grandchildren (age ranging from 6 to 21 years old at the time of sample collection) of patient 06-1 were also included for mutation screening.

The two FAP-diagnosed patients with no available information on family background were designated as 0X-1 and 0X-2, respectively. Patient 0X-1 is a 24-year old Malay male reported to suffer from altered bowel habit for the past two years and have been diagnosed with asymptomatic anaemia. On further colonoscopic examination, he was found to have developed multiple polyps along the whole length of the colon and rectum which is characteristic of FAP. Patient 0X-2 on the other hand is a female of Chinese ethnicity diagnosed with FAP at the age of 30. However, no clinical information was available for this patient.

Ethical approval for mutation studies on the subjects was obtained by relevant authorities at the Selayang Hospital. Conveyance of any of the outcome from this study is therefore at the full discretion of the collaborating physician.

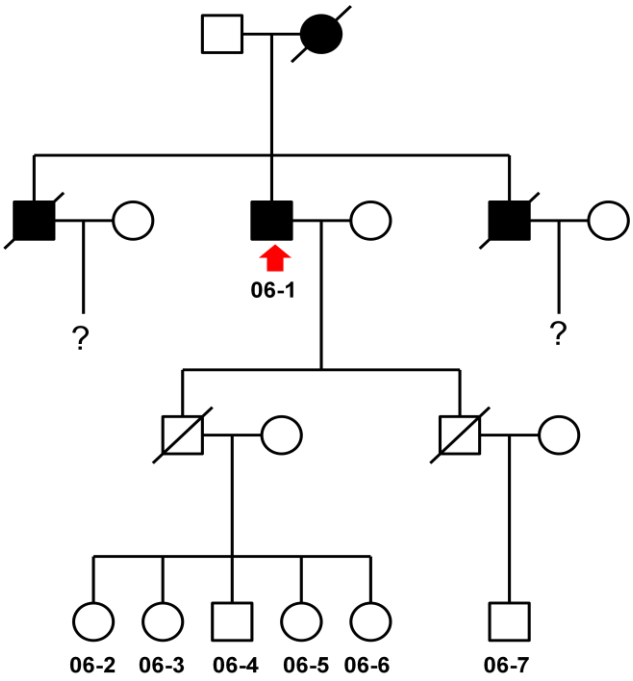

Figure 3.1: The pedigree of family FAP-06. The patient diagnosed with FAP (06-1) is indicated by the red arrow.

3.2: Genomic DNA extraction

DNA was extracted from the blood samples using QIAGEN DNA extraction kit (QIAGEN, Germany) following the instruction manual provided by the manufacturer. Prior to extraction, blood samples and RNase A solution stored in -20°C were first equilibrated to ambient temperature. To start, 20µl of Proteinase K solution (provided with the kit) were transferred to the bottom of a sterile 1.5ml microcentrifuge tube followed by 200µl of blood. Two microgram per microlitre of RNase solution was also added into the blood mixture to eliminate any trace of RNA. Equal volume of Buffer AL to the blood was then dispensed into the mixture and it was homogenised by pulse-

vortexing for approximately 15 seconds. The mixture was then incubated at 56°C for 10 minutes in a heating block to facilitate cell lysis.

After the incubation step was completed, the tube was removed from the heat block and the content of the tube was spun down briefly. Next, one volume of 100% ethanol was added to the lysed sample. The mixture was pulse-vortexed for 15 seconds, centrifuged for another 15 seconds at 12,000x*g* and transferred into a fresh spin column (provided with the kit). The transfer has to be done carefully so as not to wet the rim of the column. The column was then centrifuged at 6,000x*g* for one minute. Later, DNA washing step was carried out by adding in 500µl of ethanol-diluted Buffer AW1 to the spin column followed by centrifugation at 6,000x*g* for one minute. After discarding the filtrate, the washing step was repeated by adding 500µl of ethanol-diluted Buffer AW2 into the same spin column before it was centrifuged at 20,000x*g* for another minute.

To avoid any risk of Buffer AW2 carryover from the previous centrifugation step, the spin column was transferred into a new collection tube and centrifuged at 20,000x*g* for another minute. Later, the column was placed into a new and labeled 1.5ml microcentrifuge tube and 200µl of Buffer AE was dispensed directly to the membrane filter of the spin column. The spin column was left to stand at room temperature for one minute before it was centrifuged at 6,000x*g* for another minute. Finally, the column was removed from the tube and discarded. The DNA elution was then assessed for its quality by gel electrophoresis and quantitated using a spectrophotometer. The DNA elution is stored at -20°C for future use.

3.3: DNA quantitation

Measurement of DNA concentration and purity was carried out using a spectrophotometer (Biophotometer Eppendorf, Germany). The appropriate measurement setting and dilution factors were first keyed in before readings were taken.

Then, diluted sample was transferred into either a disposable plastic or quartz cuvette and placed into the photometer. The sample was measured thrice and averaged. Measurements of the DNA concentration, purity or OD_{600} readings were recorded for future reference.

3.4: Gel electrophoresis

Gel electrophoresis was done using agarose gel as the separation medium. The following preparation was done using electrophoresis system from Major Science (USA).

3.4.1: Gel preparation

To prepare the separation gel, agarose powder was first weighed to the appropriate amount corresponding to the desired working concentration in a dedicated conical flask. Then 1X TBE electrophoresis buffer was added to the powder to the desired volume of the gel. The mixture was boiled using a microwave for approximately 60 to 90 seconds to dissolve the agarose powder.

After the powder was completely dissolved in the buffer, the mixture was cooled down under running tap water to around 50°C before 0.1% of ethidium bromide solution was added. The mixture was swirled gently to mix the ethidium bromide and gel mixture was poured carefully into the gel caster assembled in advance. The gel caster was fixed with a comb that has the desired size and number of tooth. The pouring process must be done as quickly as possible to avoid the gel from solidifying in the flask. The gel was left to set for about 15 to 20 minutes at room temperature.

3.4.2: Electrophoresis

The hardened gel slab was removed from the cast and placed in the electrophoresis tank filled with 1X TBE electrophoresis buffer. A 5µl aliquot of the DNA sample was mixed with 1µl of 6X loading dye (30% glycerol, 20mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol) before it was loaded into a well in the gel. Then, 1µl of DNA marker (Seegene, Korea) of an appropriate indicative size was placed in a well next to the samples to facilitate the interpretation of band size.

The electrophoresis tank was connected properly to a power pack (BioRad, USA) after the sample loading was completed. The samples were then electrophoresed at 120V until the tracker dye reaches the bottom of the gel. The gel was then viewed under UV illumination (Alpha Innotech, USA) for result verification and documentation.

3.5: Polymerase Chain Reaction (PCR)

3.5.1: Primer design

Every exon of the *APC* gene was amplified using exon-specific primer pairs. The primer sequences used were based on Groden *et al.* (1991) which included a set of overlapping internal primers that were designed to amplify exon 15 in fragments of 200-500bp long (Figure 3.2). All the primer pairs were designed to generate PCR amplicons ranging from 150-500bp for SSCP analysis. The detailed description of all the primers can be found in Appendix B.

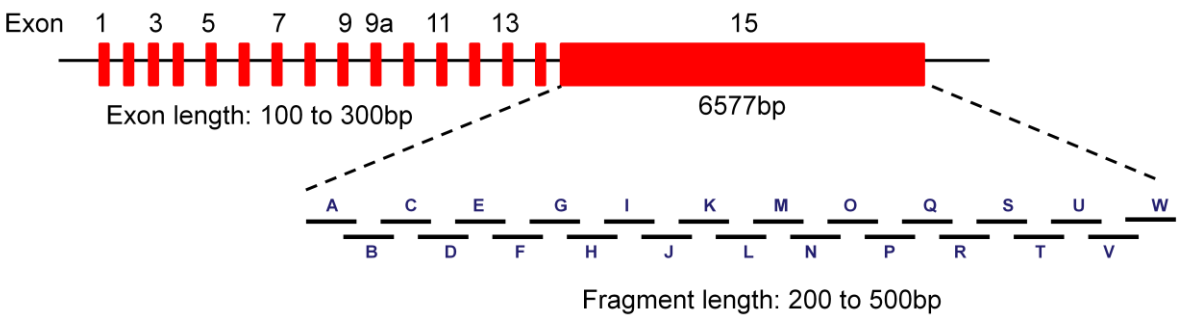

Figure 3.2: A schematic representation of PCR fragments of *APC* exon 15.

3.5.2: PCR mastermix preparation

PCR reagents were kept at -20°C for long-term storage. Before starting, the PCR reagents except for *Taq* polymerase (EURx, Poland) were thawed to ambient temperature, mixed by vortexing and spun down briefly to collect the contents to the bottom of the tube. In the meantime, the volumes of PCR reagents per reaction multiplied to the desired number of reactions were calculated.

In brief, the mastermix cocktail was prepared by adding the following reagents to the final volume of 25µl; 1X PCR buffer, 1.5mM MgCl, 200µM of each dNTP, 0.4µM of each forward and reverse primer and 1U *Taq* polymerase. Sterile distilled water was added to make the final volume up to 25µl. The mixture was then mixed and aliquoted into labeled 0.2ml PCR tube. Then, approximately 100ng of purified DNA was pipetted into each PCR tube. Additional reactions were included in the mastermix calculation and prepared to serve as negative control. The complete recipe for PCR mastermix preparation can be found in Appendix A.

3.5.3: Amplification by PCR

The PCR reactions were first mixed and briefly centrifuged after the DNA was added. The tubes were then placed in a thermocycler and PCR was carried out using a profile consisting of one denaturation step at 95°C for 5 minutes followed by 35 cycles of 30 seconds of denaturation at 95°C, 30 seconds of primer annealing at 58°C and one minute of extension at 72°C. This was followed by a five minutes final extension step at 72°C and an incubation step at 25°C for another 5 minutes. PCR products were then retrieved and either analysed immediately by gel electrophoresis as described in section 3.4 or stored at 4°C for later use.

3.6: Single strand conformation polymorphism (SSCP) analysis

3.6.1: Polyacrylamide Gel Electrophoresis (PAGE)

SSCP technique involves the use of non-denaturing PAGE to separate single strand DNA according to its conformation. In this step, non-denaturing PAGE was carried out using Protean II xi Cell electrophoretic system (Biorad, USA). The preparation of SSCP gel and electrophoresis was done following the manufacturer's instruction manual as described here. Depending on the sample number, this electrophoresis system allows the preparation and simultaneous run of a maximum of two gel sets at one time. The detailed description of all reagents required for SSCP can be found in Appendix A.

3.6.1.1: Gel cast assembly

Two clear glass plates (inner and outer plate) of different dimensions were used to cast a gel. Both glass plates were first wiped clean with 70% ethanol and subsequently 100% ethanol. Then the inner (16cm x 20cm) glass plate was treated with 200µl of Repel-Silane ES (2% dimethyldichloro-silane in octamethyl cyclotetrasiloxane) (PlusOne, USA) by wiping the glass surface evenly with the solution using SCOTT c-fold towels (Kimberly Clarks, USA). This step was repeated three times and the plate was left to dry. Repel-Silane application will prevent polyacrylamide gel from sticking to the glass plate surface.

The outer (18.3cm x 20cm) glass plate was treated with Bind-Silane (gamma-methacryloxy-propyl-trimethoxysilane) (PlusOne, USA), a chemical that can covalently attach polyacrylamide gel to a surface. This step will facilitate the staining process which involves multiple rinsing steps. The solution was distributed evenly onto the glass surface using SCOTT c-fold towels (Kimberly Clarks, USA) three times to ensure sufficient coating effect. After the surface of both plates had dried, two 0.5mm thick

spacer strips were placed along the longer edges of the outer plate's treated surface. Then the inner plate was placed on top of the outer plate with both treated surfaces facing each other. The glass plate sandwich was aligned properly before it was held in place using the sandwich clamps. The assemblage was then placed in the alignment slot of the casting stand.

3.6.1.2: Gel casting

The polyacrylamide gel for SSCP was prepared using Mutation Detetction Enhancement® (MDE) gel solution (Lonza, Switzerland) to the concentration of 0.5X as recommended by the manufacturer. The gel was prepared by mixing the following components in a beaker to a final volume of 30ml; 0.6X of TBE buffer, 0.5X of MDE gel solution, 0.1% of freshly prepared ammonium persulphate (APS) (Promega,USA), 0.1% (v/v) of N,N,N ,N -tetramethyl-ethylenediamine (TEMED) (Amresco, USA) and sufficient volume of deionized water. The addition of APS and TEMED will initiate the polymerization of the polyacrylamide-like molecules in the MDE solution. Thus, both reagents were added only when the gel is ready to be casted.

A comb of the appropriate size and number of tooth was first slotted into the space between the glass sandwich to form the sample wells. Then the homogeneous gel solution was carefully transferred into the glass plate sandwich using a micropipette until it reaches the bottom edge of the comb teeth. The comb was tilted on one side slightly before resuming with the gel pouring. This would help to release the air from the gaps in between the teeth. The comb was adjusted back to its original position when all the teeth have been covered by the gel solution. The gel was left to polymerize at room temperature for at least 60 minutes.

3.6.1.3: Sample preparation

The PCR products with the correct amplicon size were prepped for SSCP analysis while waiting for the gel to solidify. Six microlitres of each PCR product were transferred into a labelled microcentrifuge tube and were thoroughly mixed with 2µl of 4X SSCP loading dye (95% formamide, 10mM sodium chloride, 0.25% bromophenol blue and 0.25% xylene cyanole). The mixture was then incubated at 95°C for 5 minutes using a dry heat block to denature the DNA strands. This was immediately followed by snap-chilling the samples on ice to prevent the DNA strands to revert back to its double-stranded form. The samples were left on ice until they are ready to be loaded into the gel.

3.6.1.4: SSCP gel electrophoresis

To start the PAGE, the polyacrylamide gel was first fixed to the electrophoresis system by sliding the gel sandwich into place with the inner plate facing the central cooling core. Then, the acrylic buffer dam was attached to the cooling core to form the upper buffer chamber. The use of the buffer dam however can be skipped if electrophoresis was done using two sets of gel sandwich. The upper buffer compartment was filled with 0.6X TBE running buffer and checked for any leakage, indicative of improper assembly of the gel sandwich.

Sufficient volume of 0.6X TBE running buffer was then poured into the lower buffer chamber until it covers the bottom end of the gel sandwich. Then, the tank cover with its electric leads was fixed to its corresponding terminal on the chamber. The gel was pre-electrophoresed at 300V for 30 minutes at 4°C. Once the process completed, the denatured DNA samples were transferred into the gel wells using a micropipette. The samples were electrophoresed until the loading dye reaches the bottom end of the gel, at either 300V for 4 hours or 160V for 16-18 hours.

3.6.2: Gel staining

The gel sandwich was removed from the cooling core of the electrophoresis system once the electrophoresis was complete. The sandwich clamps were loosen and detached from the glass plates. The outer glass plate was removed by first pulling out one of the spacers. A spatula was then slowly pushed into the gap and twisted to pry the glass plate open. This step was done carefully to avoid breaking the gel.

The gel was then placed in a shallow tray and submerged with 1000ml of fixing solution (10% ethanol, 5% glacial acetic acid). The gel was gently agitated for 30 minutes at room temperature on an orbital shaker. After that, the fixing solution was poured into a container and will be used to stop the staining process later on.

The gel was then rinsed with ice-cold distilled water and agitated for two minutes. This step was repeated three times and followed by the staining step with 1000ml of 0.1% w/v silver nitrate solution. Staining was done with gentle agitation for one hour. The silver staining solution was then discarded and the gel was rinsed with ice-cold distilled water again to remove the excess staining solution. After that, the gel was agitated in 800ml of developing solution (0.3% formaldehyde, 3% w/v sodium hydroxide) until bands were visible with adequate contrast from the background. The developing step was stopped by soaking the gel in the used fixing solution for five minutes. The bands were then checked for any shift in mobility pattern and documented using a gel documentation system (Alpha Innotech, USA). The gel can then be discarded by scrapping it off the outer plate. Both glass plates and the casting apparatuses were cleaned from gel residues using mild detergent and air dried before storage.

3.7: PCR product clean-up

Samples with mobility shifts from the SSCP assay were purified from any PCR artifacts before they were sent for DNA sequencing. Purification was carried out using either QIAquick PCR purification kit or QIAquick gel extraction kit (QIAGEN, Germany), following the recommended procedures by the supplier.

3.7.1: Direct DNA purification using QIAquick PCR purification kit

PCR products of the same sample replicates were first pooled together into a single tube. After that, 5 volumes of Buffer PBI to the volume of pooled PCR samples were added. Then, the mixture was briefly vortexed while ensuring that the original color of the solution is retained. The color yellow indicates that the pH value of the solution is at 7.0. Ten microlitres of 3M sodium acetate (pH 5.0) can be added to the mixture to revert any change of color back to yellow.

The mixture was then transferred into a labeled spin column placed in a 2ml collection tube. The column was centrifuged for one minute to bind the DNA to the membrane. The filtrate was discarded and 750µl of 35% guanidine hydrochloride solution was then added to the column followed by centrifugation for another minute. This washing step with guanidine hydrochloride was carried out specifically to get rid of any primer dimers that might interfere with subsequent DNA sequencing step. The DNA was later washed with 750µl of ethanol-diluted Buffer PE followed by a one-minute centrifugation. Any filtrate formed was discarded and the column was centrifuged again for another minute to ensure the complete removal of washing buffer from the column.

Later, the column was transferred into a sterile 1.5ml centrifugation tube. Thirty to fifty microlitres of Buffer EB was carefully dispensed directly onto the center of the membrane. The column was left to stand at room temperature for one minute and

centrifuged for another one minute to elute the DNA. The eluted DNA was checked for its quality by gel electrophoresis as described in section 3.4 and stored at -20°C for future use.

3.7.2: DNA purification using QIAquick gel extraction kit

PCR products contaminated with unspecific amplicons can be purified by first separating them by electrophoresis on a 1% agarose gel for 25 minutes at 120V, as described in section 3.4. Amplified PCR products and its replicates were pooled into a microcentrifuge tube before they were loaded onto the agarose gel. After electrophoresis was completed, the gel was visualised using a UV transilluminator (Vilber Lourmat, France).

During this process, the band with the desired size was identified and immediately excised from the gel slab using a clean scalpel. The gel excise was trimmed from excess gel and the product was then placed in a labeled and weighed microcentrifuge tube. It is noteworthy that the gel excision step was carried out as quickly as possible to minimize UV exposure on the handler and the PCR products.

The tube containing the excised gel was again weighed and from there the gel weight was calculated. Five volumes of Buffer QG per one volume of gel weight were dispensed into the tube containing the gel. The mixture was then incubated at 50°C for 10 minutes using a heat block. The mixture was occasionally vortexed to facilitate proper dissolution of the gel as well as to ensure that the colour of the buffer stays yellow. If a colour change was observed, 10µl of 3M sodium acetate (pH 5.0) can be added to the mixture to reverse the effect.

A volume of 100% isopropanol equal to the gel weight was added into the tube after the incubation step. The mixture was homogenised and then transferred into a labeled spin column placed in a 2ml collection tube. This was followed by a centrifugation step at maximum speed for one minute to bind the DNA to the column

membrane. The filtrate was discarded and 500µl of Buffer QG was added into the column. The content was then centrifuged at maximum speed for another minute to get rid of any residual gel content.

After discarding the flow-through, the washing step was carried out by dispensing 750µl of ethanol-diluted Buffer PE into the column followed by a centrifugation step as done previously. The collection tube was emptied from any filtrate before the column was again spun at maximum speed. This repeated centrifugation step was done to ensure complete removal of any possible ethanol residue from the sample.

The spin column was then placed into a labeled 1.5ml microcentrifuge tube. Thirty to fifty microlitres of EB buffer was dispensed directly onto the center of the membrane before the column was left to stand for 5 minutes at ambient temperature. The column was then centrifuged at maximum speed for one minute to elute the DNA from the membrane. After the spinning step, the column was discarded and the eluted DNA was stored at -20°C for future use. The quantity of the purified DNA was later checked by gel electrophoresis as described in section 3.4.

3.8: DNA sequencing

Purified samples were sequenced using 3130*xl* Genetic Analyzer from Applied Biosystem (ABI, USA). Sequencing services were provided by Center for Research in Biotechnology for Agriculture (CEBAR), University of Malaya. The sequencing results were then analysed using Applied Biosystem's Sequence Analysis software v5.3 with KB basecaller v1.4 and Chromas version 4.0 (Technelysium Pty Ltd).

3.9: Splicing assay

In vitro gene expression approach was opted to assess the effect of c.847C>T mutation on exon 8 splicing regulation. In sum, exon 8 of the *APC* gene was first

amplified and cloned into a cloning vector before it was engineered to carry the c.847C>T mutation by site-directed mutagenesis. Following that, a minigene was constructed by joining PCR-amplified exon 7, exon 8 that carries c.847C>T mutation and exon 9 of the *APC* gene using overlap extension PCR method. The construct was then cloned into a mammalian expression vector and transfected into cultured mammalian cancer cells for *in vitro* expression. The mRNA transcripts produced by the transfected cells were purified and analysed for any variation in the size of the transcripts. The following subsections will explain in detail the procedures undertaken.

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

Exon 7, exon 8 and exon 9 were amplified using primers published by Groden *et al.* (1991) as described in section 3.5.1 and following the steps described in section 3.5. The PCR-amplified exons fragments were purified using the purification method described in section 3.7.

3.9.2: DNA cloning in bacterial hosts *Escherichia coli (E. coli)*.

3.9.2.1: Preparation of bacterial growth broth and plated media

Luria Bertani (LB) growth media were used to grow bacteria cells in this DNA cloning step. Preparation of LB broth was done by first dissolving 20g of LB media powder (Pronadisa, Spain) in 900ml of distilled water in a clean beaker by using a magnetic stirrer. Then, the mixture was added with an appropriate volume of distilled water to make the final volume of 1000ml. Then, 10ml aliquots of the media broth were transferred into clean universal bottles and autoclaved for 20 minutes (121°C, 151lb/sq). Autoclaved LB broth was kept at room temperature for future use.

LB agar plates were prepared by first dissolving 14g of LB agar powder mix (Pronadisa, Spain) in 300ml of distilled water in a sterile Schott bottle using a magnetic

stirrer. Then, the mixture was added with sufficient volume of distilled water to make the final volume of 400ml. The media was then sterilised by autoclaving at the same setting as described previously. After the sterilisation step was completed, the LB media was cooled off under running tap water until the temperature reaches approximately 50°C. Then, 50µg/ml of ampicillin, 80µg/ml of X-Gal and 0.5mM of IPTG were added into the media before it was mixed by gentle swirling. The LB media was then poured into labeled disposable plastic Petri dishes until approximately half-full and left to solidify at the ambient temperature for about two hours. The LB agar media that have hardened were kept in 4°C for future use. Both LB broth and agar media were prepared fresh before being used in subsequent competent cell preparation, transformation and plasmid isolation steps.

3.9.2.2: Preparation of competent cells

E. coli competent cells were prepared by following the Rubidium chloride competent cell preparation method. A detailed description for the preparation of reagents used in this step can be found in Appendix A. A 2µl aliquot taken from JM109 bacteria strain stock (Promega, USA) was inoculated into 10ml of LB broth and cultured at 37°C with constant agitation at 250rpm for more than 16 hours. Then the cells were subcultured in another 10ml of LB broth at the same culture settings and grown to log phase. The subculture cell density was measured every 5 minutes until it reaches OD_{600} value of 0.5 which is indicative of log phase of growth. Next, the subculture was transferred into a prechilled 15ml tube and the cells were pelleted by centrifugation at 4°C for 5 minutes at the speed of 1060x*g*.

The supernatant formed was carefully decanted before 5ml of ice-cold RF1 solution was added into the tube. The pellet was then resuspended by pulsed inversion on ice. The suspension was later incubated on ice for 20 minutes and centrifuged for 15

minutes to pellet the cells as previously described. The supernatant formed was then carefully decanted. The pellet was resuspended with 200µl ice-cold RF2 solution by gentle inversion on ice. The competent cell suspension was then transferred into prechilled 1.5ml microcentrifuge tubes, snap-frozen in liquid nitrogen for five seconds and stored at -80°C for future use in bacterial transformation.

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

The PCR-amplified products from section 3.9.1 were cloned into pGEM-T Easy cloning vector (Promega, USA) to serve as glycerol stocks for future use. Cloning of exon 8 is also necessary to facilitate the site-directed mutagenesis step later on. Briefly, the DNA of interest was first ligated to the plasmid vector before the vector was transformed into competent bacterial cells. The bacteria were then grown on a growth media overnight and screened for positive transformant the next day. All the aforementioned DNA cloning steps were carried out according to the protocol provided by the manufacturer and are described in details in the following subsections.

3.9.2.3a: Ligation

A ligation reaction was set up according to the recommendations in the instruction manual supplied with the kit. The pGEM-T Easy vector and 2X Rapid ligation buffer were thawed on ice and briefly centrifuged prior to use. The ligation reaction was prepared by mixing 1X of Rapid Ligation buffer, 50ng of pGEM-T Easy vector and 150ng of purified exon 8 DNA in a 0.5ml microcentrifuge tube. The concentration of the DNA used in the reaction was three times the concentration of the vector, but the ratio may be optimised by increasing the DNA content up to 8 times the vector concentration. Three Weiss units of T4 DNA ligase and sufficient volume of sterile distilled water were then added to the mixture to top the reaction volume up to 10µl.

The reaction was spun briefly to collect any droplets on the tube wall to the bottom of the tube and then it was incubated at 4°C for 16-18 hours.

3.9.2.3b: Transformation

Five microlitres of the ligation product were transferred into a prechilled 1.5ml microcentrifuge tube containing 100µl of JM109 competent bacterial cells (Promega, USA). The mixture was then incubated on ice for 30 minutes. The bacteria mixture was heat shocked at 42°C for 45 seconds and snap-chilled on ice immediately for another two minutes to let the cells recover from the heat stress. The bacterial mixture was added with 900µl of LB growth media and mixed briefly by pipetting. The mixture was then incubated at 37°C with medium speed agitation in a hybridization oven (Shel Lab, USA) for three hours to allow the cells to grow.

The bacterial culture was retrieved from the oven after the incubation and was spun down at 100x*g* for 10 minutes to concentrate the cell to the bottom of the tube. About 2/3 of the supernatant was discarded using a pipette before the cells were resuspended by gentle pipetting in the remaining supernatant. About 100µl of the cell suspension was then transferred onto a dry LB agar plate and was spread evenly on the agar surface using a sterile glass rod spreader. This step was carried out in an aseptic condition to avoid contamination by other microorganisms. The plate was then placed in an incubation oven set at 37°C to allow the bacteria to grow overnight. The remainder of the bacterial culture was stored at 4°C if re-plating is required later on.

3.9.2.3c: Colony selection by PCR

The use of pGEM-T Easy vector enables identification of transformation-positive bacteria colonies to be done using blue-white colony selection method. A single white colony was gently lifted from the LB agar surface using either a sterile wire loop or a

toothpick. A portion of the isolated colony was first transferred onto a 6 x 6 gridded LB agar plate to serve as a library plate. The rest of the colony portion was then mixed with 30µl of sterile distilled water in a 0.5ml microcentrifuge tube. These steps were repeated with the other positive colonies until the required number of isolates was obtained. These bacterial colony suspensions were then boiled in a heating block for about 10 minutes and briefly centrifuged to collect the suspension to the bottom of the tube.

The denatured cell suspensions were then used as DNA templates in a PCR-based assay to screen for colonies that carry the correct exon 8 insert. DNA fragments inserted in the cloning site would be simultaneously amplified when M13 forward and reverse primers were used to amplify the corresponding portion of the vector. This PCR screening step was carried out as described in section 3.5 with the annealing temperature set at 60°C for optimal M13 primer binding. The PCR products were then analysed by agarose gel electrophoresis as described in section 3.4 and colonies carrying the desired inserts were identified based on the presence of DNA bands that corresponded to the length of the inserted fragment.

3.9.3: Preparation of plasmid DNA by alkaline lysis with SDS

Bacterial colonies that have been positively identified to carry the correct inserts were isolated from the library plate and inoculated into 10ml of LB broth media supplied with 50µg/ml of ampicillin. The inoculants were cultured at 37°C for 16-18 hours with agitation at 250rpm. A 900µl aliquot of the cell suspension was transferred into a labeled 1.5ml microcentrifuge tube containing 100µl of glycerol once overnight culture was complete. The mixture was mixed by pipetting and then stored at -80°C as glycerol stock. The rest of the cell suspension was transferred into a labeled 15ml tube and centrifuged at 4260x*g* for 15 minutes.

The supernatant formed was decanted entirely without disturbing the pellet. Then 200µl of alkaline lysis solution I (50mM glucose, 10mM EDTA, 25mM Tris-Cl) were pipetted to the cell pellet. The pellet was homogenised by vigorous vortexing and transferred into a sterile and labelled 1.5ml microcentrifuge tube before 200µl of alkaline lysis solution II (0.2N NaOH, 1% SDS) were added. The tube was gently inverted several times to homogenise the mixture. Then, it was left to stand at room temperature for 4 minutes followed by the addition of 200µl of alkaline lysis solution III (3M KoAC in acetic acid). The mixture was then incubated on ice for 15 minutes before it was centrifuged at 15,700 x*g* for 10 minutes. The supernatant was aspirated out from the tube as much as possible and transferred into a new sterile 1.5ml microcentrifuge tube while avoiding carry over of the pelleted cell debris. Three microlitres of 50µg/µl RNase A solution was pipetted into the supernatant and the mixture was incubated at 37°C for three hours to eliminate any trace of RNA contamination.

Six hundred microlitres of Tris-saturated phenol was added to the mixture after the RNase treatment and the mixture was briefly vortexed. The mixture was then centrifuged at 15,700x*g* for three minutes causing the mixture to separate into two distinct layers. The aqueous upper layer was carefully aspirated out from the tube to avoid carry over of the intermediate white layer and transferred into a fresh sterile 1.5ml microcentrifuge tube. Then 600µl of chloroform were added to the solution and again mixed by vortexing. The mixture was later centrifuged at 15,700x*g* for three minutes. The aqueous upper layer formed after this centrifugation step was collected as much as possible without disturbing the bottom lower layer and transferred to another fresh sterile 1.5ml microcentrifuge tube.

Two volumes of isopropanol and 0.1 volume of 5M sodium chloride solution were then added to the aqueous solution based on the volume of the aqueous solution recovered from the chloroform washing step. The mixture was incubated on ice for 30

to 60 minutes to precipitate the DNA followed by centrifugation at 15,700x*g* for 20 minutes to pellet the precipitated DNA. The supernatant was then removed by gentle aspiration to avoid accidental pellet removal. One millilitre of 70% ethanol was then added to the DNA pellet and the tube was spun again at the same setting for five minutes. The supernatant was removed as much as possible by aspiration without disturbing the pellet.

The pellet was then dried by centrifugation under vacuum condition for five minutes using SpeedVac Concentrator (HETO, Denmark). The tube was removed from the machine with care and the dried pellet was reconstituted by adding 50µl of sterile distilled water to it. The plasmid sample was stored at -20°C for future use. The plasmid insert was then sequenced as described in section 3.8 to verify the insert sequence.

3.9.4: Site-directed mutagenesis

The mutation of interest, c.847C>T was introduced to the normal exon 8 sequence by site-directed mutagenesis method using QuikChange Site-Directed Mutagenesis Kit (Stratagene, USA) as shown in Figure 3.3. The instruction manual provided was referred to when using the kit. A pair of new primers (SDM-F and SDM-R; see Appendix B) that contain the desired mutation and can anneal to the same sequence on the opposite strands of the plasmid was also designed according to the manufacturer's guidelines. This primer pair was commercially synthesised and later used in the mutant strand (8m) synthesis by thermal cycling. The isolated plasmid carrying the normal exon 8 sequence was used as the template for the mutagenesis step.

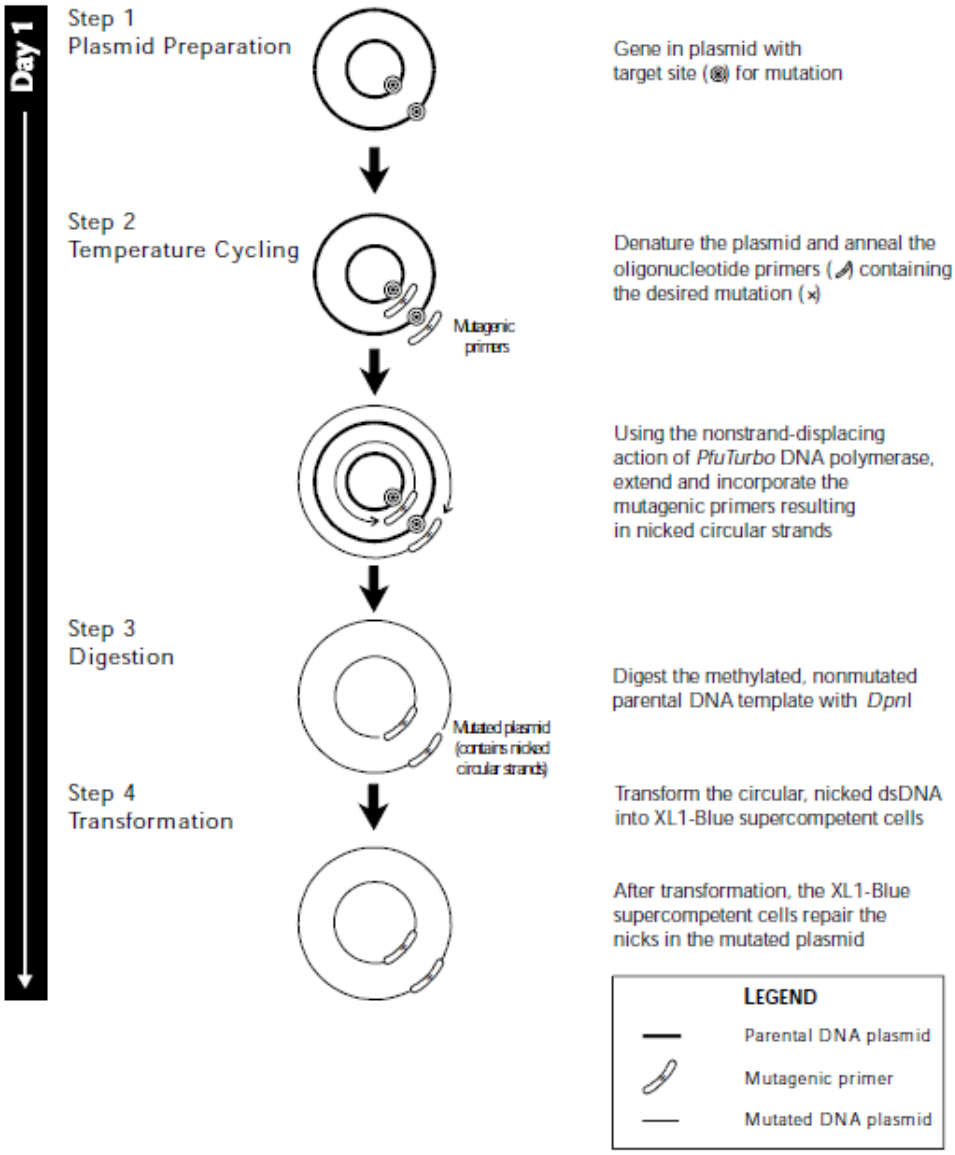

Figure 3.3: A schematic representation of site-directed mutagenesis.
(Source: QuikChange Site-Directed Mutagenesis Kit Handbook, 2009)

3.9.4.1: Mutant strand synthesis reaction preparation

Two sample reactions using 10ng and 50ng DNA templates each were prepared by mixing the following reagents in a thin-walled tube: 1X reaction buffer, 10 or 50ng double stranded DNA template, 125 ng of both SDM-F and SDM-R primers, 1µl dNTP mix, 2.5U *PfuTurbo* DNA polymerase and sufficient volume of sterile distilled water to make a final volume of 50µl. It is noteworthy that the DNA polymerase was added last to the reaction as recommended by the protocol.

Another reaction was also prepared to serve as a control for the mutagenesis step. The reaction consists of 1X reaction buffer, 10ng of pWhitescript 4.5-kb control plasmid, 125 ng of both control primers #1 and #2, 1µl of dNTP mix, 2.5U of *PfuTurbo* DNA polymerase and sufficient volume of sterile distilled water to make a final volume of 50µl.

3.9.4.2: Thermal cycling parameter

The mutant strand synthesis reactions prepared were placed in a thermocycler equipped with a hot-top assembly. The reactions were heated to 95°C for 30 seconds then to 18 cycles of the following temperature cycling program: 95°C for 30 seconds, 55°C for one minute and 68°C for five minutes. After the cycles were completed, the reactions were placed on ice for two minutes to cool the reaction to ≤37°C. Then 1µl of *Dpn*I restriction enzyme (10U/µl) was added to each of the amplification reaction. The mixture was mixed gently by pipetting and later briefly centrifuged. The reactions were then incubated at 37°C for one hour to digest the methylated non-mutated parental DNA template.

3.9.4.3: Transformation of XL1-Blue supercompetent cells

The newly-synthesised plasmids that were not digested by *Dpn*I were transformed into XL1-Blue supercompetent cells according to the steps described in the instruction manual. Briefly, the competent cells were first thawed on ice before 50µl of the competent cells were aliquoted to prechilled 1.5ml microcentrifuge tubes for each reaction required. Then 1µl of each *Dpn*I-treated DNA sample were added to the competent cells aliquots. The transformation reaction were swirled gently to mix and then incubated on ice for 30 minutes.

The reactions were later heat-shocked for 45 seconds at 42°C and immediately

snap-chilled on ice for two minutes. This was followed by the addition of 500µl of NZY^+ broth preheated at 42°C to each reaction. The bacterial suspensions were then incubated at 37°C for one hour with 225-250rpm agitation. Then 250µl of each reaction was plated on LB agar plates containing 100mg/ml of ampicillin, 80µg/ml of X-Gal and 20mM of IPTG. The plates were incubated at 37°C for at least 16 hours.

3.10.4.4: Colony selection

Similar to the steps described in section 3.9.2.3c, ten colonies from each DNA concentration tested that grew on the selection plate were isolated and transferred to a library plate and at the same time inoculated in 10ml of NZY^+ broth for plasmid isolation. Plasmid isolation was done following the steps described in section 3.9.3 and the purified plasmids were subjected to colony PCR to verify the presence of the insert of interest. PCR was done as described in section 3.5 using intronic forward and reverse primers of exon 8. The insert was also sequenced to ensure that the c.847C>T mutation was successfully generated without any additional mutations elsewhere in the insert sequence.

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

3.9.5.1: Primer design

Intronic primers were used to amplify exon 7, exon 8 and exon 9 individually. Each primer sequence was designed to bind approximately 30 base pairs upstream and downstream from the intron-exon junction of the exon to preserve the splicing regulatory motifs that are located within these regions (Figure 3.4; panel A). Ex 7R - Ex 8F primer pair and Ex 8R - Ex 9F primer pair were designed to carry twelve nucleotide overlapping region at their 5’ ends to facilitate minigene construction by overlap extension PCR (SOE-PCR) method described in section 3.9.5.3. Internal primers

labeled Ex 9IntF and Ex 9IntR were also designed to facilitate exon 9 splicing to the rest of the construct. At the same time, the forward primer of exon 7 (Ex 7R) and the reverse primer of exon 9 (Ex 9R) were also designed to carry an *Eco*RI cut site at the 5' end of the primer sequence. This addition was needed to facilitate ligation of the minigene into the pTARGET mammalian expression vector.

Sequence Manipulation Suite version 2 software (Stothard, 2000) was used to describe the properties of each primer sequence including its molecular weight, melting temperature, percentage of GC content as well as its PCR suitability. The best primers were selected and commercially synthesised before they were tested in a trial PCR run to verify their functionality and specificity.

3.9.5.2: PCR amplification of the exons using overlap primers

Exon 7, exon 8, fragment 9a and fragment 9b were amplified using primers designed in section 3.9.5.1 using DNA templates from plasmid extractions described in section 3.9.3 for exon 7, exon 8 and exon 9 as well as in section 3.9.4.4 for exon 8m respectively (Figure 3.4; panel A). The PCR-amplified exon fragments were purified using the purification method described in section 3.7 before being used in SOE-PCR.

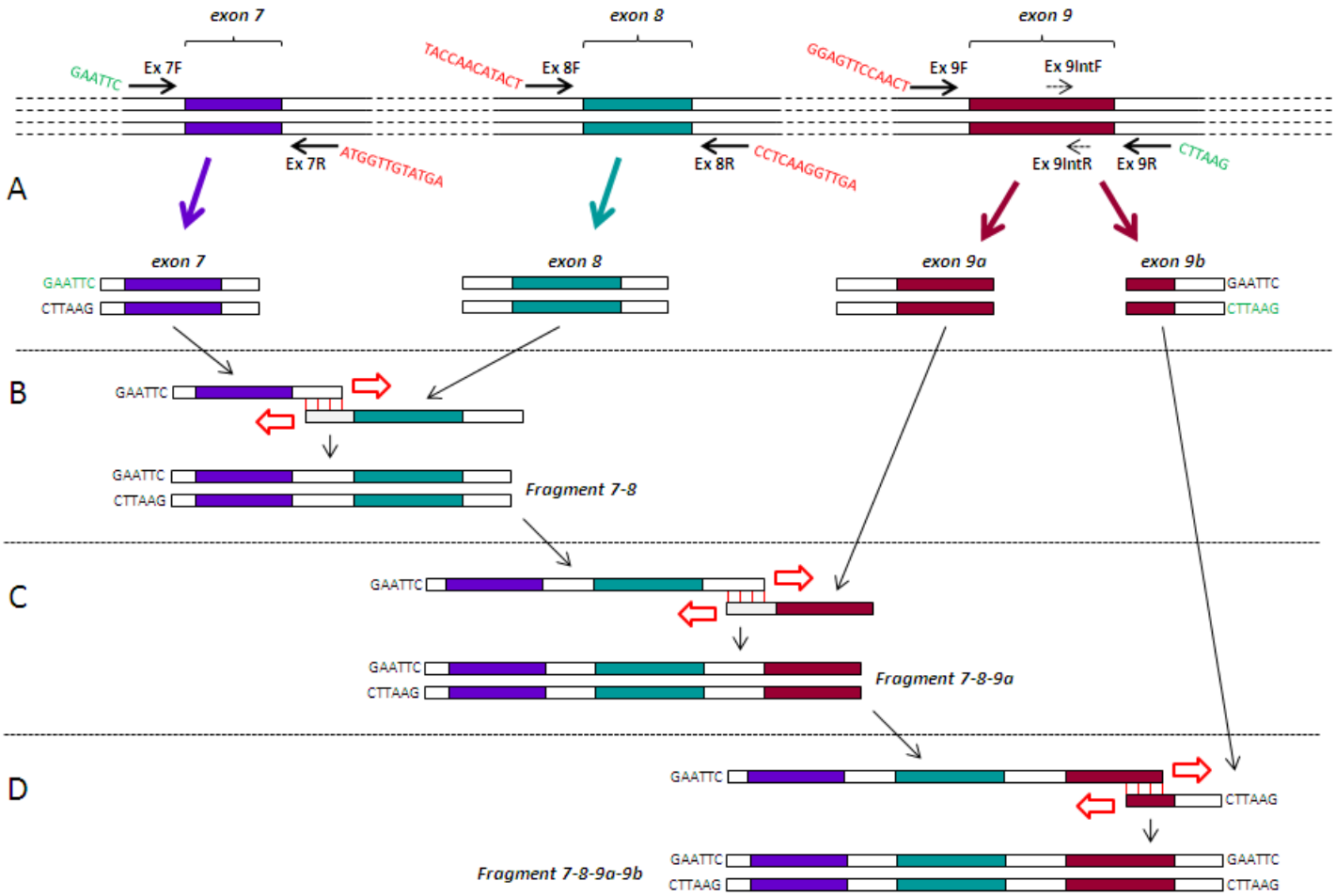

Figure 3.4: A schematic representation of SOE-PCR. **(A)** Individual exons were first amplified using exon-specific primers either carrying an *Eco*RI motif (green) or 12bp overlap sequences (red). **(B)** Exon 7 was joined to exon 8 to produce fragment 7-8. **(C)** Fragment 7-8 was joined to exon 9a to produce Fragment 7-8-9a. **(D)** Fragment 7-8-9a was joined to exon 9b to produce the complete minigene (fragment 7-8-9a-9b). Red arrows: DNA synthesis direction using free 3' ends as start points for strand extension.

3.9.5.3: SOE-PCR

Two types of minigenes that consisted of exon 7, exon 8 and exon 9 of the *APC* gene were constructed using this method. The difference between the two constructs is one minigene has the native *APC* exon 7, exon 8 and exon 9 sequences (designated as construct 7-8-9) while the other contains exon 8 that carries the c.847C>T mutation engineered previously as described in section 3.9.4 (designated as construct 7-8m-9). Figure 3.4(B-D) shows the schematic representation of the minigene construction by SOE-PCR.

SOE-PCR was carried out in two steps; the first PCR was to overlap and fuse two DNA fragments together while the second PCR was carried out to selectively amplify the overlapped products. In the first step, the purified exon 7 was mixed with either the native (exon 8) or the mutant exon 8 (exon 8m) in a single reaction. The mastermix properties and the PCR program used were the same as described in section 3.5.2 but the number of PCR cycle was reduced to only 10 cycles. The joined DNA fragment from this step was then diluted 10-fold before it was used as template for the second step of SOE-PCR. The second SOE-PCR step was done according to the same condition of the first step SOE-PCR except the reaction was continued for 32 cycles. The amplification products (designated as fragment 7-8) of the second step of SOE-PCR were then viewed under UV light as described previously for size verification.

Both steps in SOE-PCR were then repeated to join fragment 7-8 to exon 9a and subsequently fragment 9b. All successfully fused exon fragments at every splicing step were purified and cloned into pGEM-T cloning vector and were kept as glycerol stock. The vector plasmids containing the DNA construct inserts were also isolated and the insert DNA sequence was verified by sequencing described in section 3.8.

3.9.6: Cloning of minigene into pTARGET mammalian expression vector

3.9.6.1: *Eco*RI digestion

Cloning of the constructed minigene into the pTARGET vector was done by first isolating the construct from the pGEM-T Easy vector using *Eco*RI restriction endonuclease (RE). At the same time, the pTARGET vector was also digested with the same enzyme to facilitate the ligation of the construct to the vector. The RE reactions were set up as follows: 1X restriction enzyme buffer, 0.1µg/µl acetylated BSA, 1µg DNA template, 1U *Eco*RI restriction enzyme and topped up with sterile distilled water to a final volume of 20µl. These reaction mixtures were briefly centrifuged before it was incubated at 37°C for three hours. After the incubation, all the reactions were electrophoresed on 1% agarose gel and then viewed under UV illumination to confirm the effectiveness of the digestion step.

3.9.6.2: Ligation

Ligation reaction was carried out to join the minigene to the pTARGET vector following the manufacturer's protocol. The reaction was prepared by mixing 1X T4 DNA Ligase buffer, 60ng pTARGET vector, 200ng purified minigene construct and 3 Weiss units of T4 DNA Ligase in a 0.5ml centrifuge tube. The mixture was topped up with sterile distilled water to a final volume of 10µl and was later incubated at 4°C for at least 16 hours.

3.9.6.3: Transformation

Transformation of the pTARGET into *E. coli* strain JM109 was carried out according to the protocol described in section 3.9.2.3b.

3.9.6.4: Colony selection

Selection of transformed colonies was carried out based on the blue/white colony screening method. The bacterial DNA preparation and the PCR-based insert verification were done as described in section 3.9.2.3c with the exception of the PCR primers used. Colony PCR was done using T7 promoter forward primer and Ex9.3 reverse primer. The vector-insert primer combination was used to selectively amplify only the insert ligated to the vector in the correct orientation.

Amplification products was separated by 1% agarose gel electrophoresis as described in section 3.4 and viewed under UV illumination. The positive colonies were identified based on the presence of DNA bands with the correct band size and proceeded with plasmid isolation step described in section 3.9.3. The vector plasmids isolated were then sequenced as described in section 3.8 using T7 promoter and pTARGET sequencing primers to verify the cloned DNA sequence.

3.9.7: Cell culture

3.9.7.1: Preparation of culture medium

Cells were cultured in Dulbecco's Modified Essential Medium (DMEM) (GIBCO, USA). The 10X DMEM base was prepared by adding 26.8g DMEM powder, 6g sodium bicarbonate and 9.56g HEPES buffer together in a sterile conical flask filled with 100ml of distilled water. The mouth of the flask was covered with a parafilm strip and the mixture was dissolved using a magnetic stirrer.

The pH of the solution was later adjusted to 7.0 using a pH meter (Sartorius, Germany). The mixture was then sterilised by filtration through 0.2µm pore size membrane filter (Nalgene, USA). This step was carried out in aseptic condition to avoid microbial contamination. The collection bottle was then capped and sealed using parafilm strips. The 10X medium stock solution was stored at 4°C.

Cells were grown in 1X DMEM growth medium [10% of 10X DMEM, 10% of inactivated fetal bovine serum (FBS; GIBCO, USA) and 1% ampicillin (GIBCO, USA)]. To prepare 100ml of the 1X DMEM growth medium, 10ml of 10X DMEM stock and 10ml of deactivated FBS were first dispensed into a sterile Schott bottle. The mixture was then added with 79ml of autoclaved ultrapure distilled water. Then, 5ml of the medium solution were aliquoted into a Petri dish and incubated overnight at 37°C in an incubator supplied with 5% CO_2 (Heraeus, USA) to act as an indicator that the media preparation is free from any microbial contamination. Ampicillin can only be added into the medium the next day if the overnight culture is aseptic. Ampicillin however can be omitted from the preparation if the use of antibiotic-free growth medium is required. FBS concentration in the medium can also be reduced to 1% v/v to slow down cell doubling and maintain the cells at a certain density. All growth and maintenance media were kept at 4°C.

3.9.7.2: Cell maintenance

Three to five days old growth media was removed from the cultured cells by pipetting and later, the cells were rinsed twice with 1ml of Dulbecco's phosphate buffer saline (PBS; Sigma, USA) to ensure the complete removal of the media. Then, 5ml of fresh growth medium were carefully dispensed into the culture flask and replaced in the 37°C incubator with 5% CO_2 to allow the cells to grow.

Cells that have reached 70-100% confluence need to be apportioned to create space for new cells to grow. The old growth medium was first removed from the culture flask. The cells were then rinsed twice with 1ml of PBS to get rid of residual growth medium. One millilitre of TrypLE (GIBCO, USA), an analogous enzyme of trypsin, was later dispensed into the culture flask and the cells were trypsinized at 37°C for 3-5 minutes. The flask was rapped against the operator's palm to facilitate disassociation of

the cells from the flask surface. Later, 2ml of 1X growth medium were directly dispensed to the cells to terminate the cell dissociation step.

The whole mixture was then transferred into a 15ml centrifuge tube and centrifuged at 393x*g* for 5 minutes (Beckman Coulter, USA) to separate the cells from the protease. After the cells were pelleted, the supernatant was decanted and the cells were resuspended in 2ml of 1X growth medium by pipetting until the cell were no longer clumped. The condition of the cells was assessed by observing them under an inverse microscope (Olympus, Japan). The cell suspension was split into two portions by adding another 8ml of 1X growth medium to the suspension and aliquoting 5ml of the suspension each into new 25cm^2 culture flasks (NUNC, Denmark). The cells were then incubated at 37°C with 5% CO_2 to allow the cell growth.

3.9.7.3: Cell cryopreservation and revival

The cells need to first be detached from the flask by trypsinization prior to cryopreservation. The trypsinised cells were then transferred into a 15ml centrifuge tube and centrifuged at 393x*g* for 5 minutes. The resulting supernatant was removed and the cell pellet was resuspended in 2ml of cryosolution (90% FBS, 10% DMSO). Another 8ml cryosolution was added after the cells were completely resuspended. The mixture was aliquoted 1ml into each cryotubes (NUNC, Denmark) and kept at -80°C overnight. Following this, the tubes were transferred into a liquid nitrogen storage tank for long term storage.

Cryopreserved cells were revived by first thawing the cells to room temperature. Then, the cell suspension was transferred into a 15ml centrifuge tube and centrifuged at 393x*g* for 5 minutes. The resulting supernatant was decanted and the pelleted cells were resuspended in 2ml of 1X growth medium. Then, another 3ml of 1X growth medium were added to the cells and the mixture was carefully transferred into a 25cm^2 culture

flask and incubated at 37°C with 5% CO_2 for cell growth. The following day, the medium was changed to remove any non-viable cells.

3.9.7.4: Cell enumeration

Cell counting was performed using a haemocytometer (Cascade Biologics, USA) as described in the manufacturer's protocol. The haemocytometer was first cleaned by wiping its surface with 70% alcohol. Then a clean glass cover slip was positioned on the ruled grids of the haemocytometer. The cell suspension was diluted 10X with trypan blue solution (Sigma, USA) to the final volume of 100µl. The mixture was then transferred onto the gridded plane of the haemocytometer by gently pipetting the mixture to the side of the slipcover. The mixture was let to seep through the gap in between the slipcover and the haemocytometer gridded plane.

The haemocytometer was observed under an inverse microscope once the mixture had completely spread over the surface. Viable cells were not stained by trypan blue while dead cells appeared blue in color. Thus, cells that appear colorless in the four 1mm x 1mm boxes in each corner of the gridded plane were counted. The concentration of the cells (cells/ml) was determined by averaging the total numbers of cells from all four counting areas and multiplying the value to the dilution factor (x 10^4).

3.9.8: Transfection

In vitro expression of the minigene construct requires the expression vector to be transported into the host cells. In this study, the vector was transfected into hepatocarcinoma cell line HepG2 using Optifect™ reagent (Invitrogen, USA). Transfection was carried out according to the manufacturer's protocol. The appropriate volume of growth medium, Optifect™ and Opti-MEM reagents for cell transfection in a 24-well plate were prepared according to the scale-up manual provided in the kit.

The transfection reaction was prepared in a sterile 1.5ml microcentrifuge tube by first diluting 2µl of Optifect™ reagent in 50µl Opti-MEM Reduced Serum Medium. The mixture was gently homogenised by pipetting and left to stand at ambient temperature for 5 minutes. In the meantime, 1µg vector plasmid was diluted in 50µl Opti-MEM Reduced Serum Medium in a separate 1.5ml microcentrifuge tube. The Optifect™-Opti-MEM mixture was added into the tube containing the plasmid DNA after the 5-minute incubation period. The mixture was then gently flicked to mix followed by a 20-minute incubation period at ambient temperature.

The entire mixture was later added to the cell culture plate containing 50-70% confluent cells. The plate was gently rocked sideways by hand to mix the media before it was incubated at 37°C in a CO_2 incubator. The growth medium containing the transfection reaction was replaced with fresh medium after 4 to 6 hours and the cells were left to grow for another 24-72 hours.

3.9.9: RNA extraction

Transfected cells were harvested 48 hours post-transfection and prepped for total RNA extraction using TRIZOL reagent (Invitrogen, USA). A rubber cell scrapper was used to dissociate the cells from the bottom of the well and the entire cell suspension was transferred into a 15ml centrifuge tube. The cells were pelleted by centrifugation at 393x*g* for 5 minutes (Beckman Coulter, USA) and the supernatant formed was discarded.

Then 1ml of TRIZOL per 3.5cm diameter plate of cells pelleted was added into the tube. The cells were gently lysed and homogenised by passing them through a pipette several times. Then the cell lysate was incubated at ambient temperature for 5 minutes before 200µl of chloroform per 1ml TRIZOL used was added to the tube. The mixture was vigorously shaken by hand for 15 seconds and left to stand for another 3

minutes at the ambient temperature. The mixture was later centrifuged at 12,000x*g* for 15 minutes at 4°C.

The aqueous phase was aspirated out and placed in a fresh 1.5ml microcentrifuge tube to which 500µl of isopropanol per 1ml TRIZOL used was later added. The mixture was incubated at ambient temperature for 10 minutes and then centrifuged for another 10 minutes using the previous centrifugation setting. The supernatant formed was discarded without disturbing the RNA pellet.

One millilitre of 75% ethanol to 1ml of TRIZOL used was added to wash the pellet before it was centrifuged at the speed of 7500x*g* for 5 minutes at 4°C. The ethanol was carefully pipetted out as much as possible before the pellet was stand to dry at room temperature. When the pellet was almost dry, 30µl RNase-free water was added passed through the pipette several times. The mixture was then incubated at 60°C for 10 minutes to facilitate reconstitution of the RNA pellet and later stored at -80°C for future use.

3.9.10: Reverse Transcriptase PCR (RT-PCR)

3.9.10.1: RT-PCR mastermix preparation

The transcription products of the transfected minigene were detected by RT-PCR using SuperScript III One-Step RT-PCR System with Platinum *Taq* DNA polymerase (Invitrogen, USA). This assay was done according to the instruction manufacturer's protocol. An RT-PCR reaction comprised of 1X Reaction Mix, 100ng template RNA, 10µM of each forward and reverse gene-specific primers, 2µl SuperScript III Reverse Transcriptase with Platinum *Taq* DNA Polymerase and RNase-free water topped-up to a final volume of 50µl was prepared in a 0.2ml thin-walled PCR tube on ice.

Two control reactions were included for every batch of RT-PCR performed. A control reaction was prepared by replacing the SuperScript III Reverse Transcriptase

with Platinum *Taq* DNA polymerase with regular *Taq* DNA polymerase to facilitate detection of DNA contamination. Another RT-PCR reaction devoid of any RNA template was also included for every RT-PCR batch as negative control.

3.9.10.2: Thermal cycling profile

The tubes containing the RT-PCR reactions were placed in a thermocycler and PCR was performed using the PCR profile as recommended in the instruction manual provided with the kit. The PCR started with 30 minutes incubation at 55°C to allow cDNA synthesis. This was immediately followed by a 2-minute denaturation step at 94°C and 40 cycles of 15 seconds DNA denaturation at 94°C, 30 seconds primer annealing step at 58°C and extension for 60 seconds at 68°C. A final extension step at 68°C for 5 minutes was added to the cycle to allow complete DNA synthesis before the reaction was terminated. Amplification products were then separated by agarose gel electrophoresis and viewed under UV illumination as described in section 3.4.

4.0: Results

4.1: PCR-SSCP assay

4.1.1: PCR amplification of *APC* gene exons

All 15 exons of the *APC* gene were successfully PCR-amplified using primer pairs as described by Groden *et al*., 1991. Figures 4.1 (a) and (b) show examples of successful PCR amplification of the *APC* gene exons.

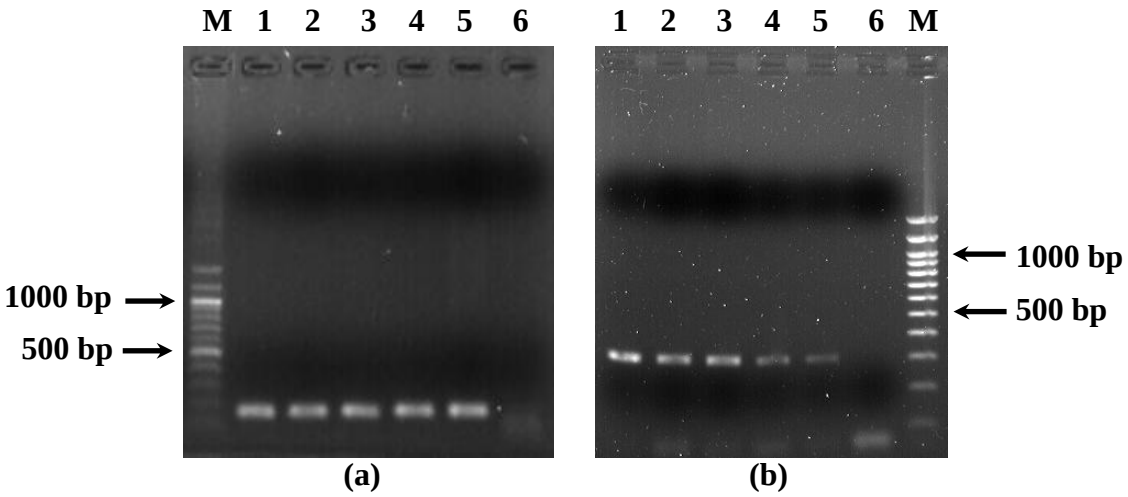

Figure 4.1 Examples of agarose gel electrophoresis of PCR products for **(a)** exon 8 (184bp) and **(b)** fragment 13 (306bp).

Lane 1: 06-1
Lane 2: 06-2
Lane 3: 0X-1
Lane 4: 0X-2
Lane 5: Unaffected control DNA
Lane 6: Negative control
Lane M: Marker (100bp ladder)

4.1.2: SSCP analysis

4.1.2.1: SSCP analysis of exon 8

SSCP results of exon 8 for samples 0X-1 and 0X-2 showed distinct band mobility shift when electrophoresed on 0.75X MDE gel at 4°C overnight (Figure 4.2). Sample 0X-1 showed the presence of an extra band slightly below the top most band when compared to the banding pattern of the unaffected control. Meanwhile sample 0X-2 was found to have an additional bright band when compared to the same unaffected control sample. Samples

06-1 and 06-2 however showed SSCP banding pattern that are consistent with the unaffected control sample.

4.1.2.2: SSCP analysis of exon 13

SSCP results of exon 13 from samples 06-1 and 06-2 exhibited variation in banding pattern when compared to an unaffected control sample as shown in Figure 4.3. Sample 06-1 showed the presence of two additional bands when compared to the four main bands found in the control. Samples 0X-1 and 0X-2 also exhibited an additional band when compared to the control. In contrast, sample 06-2 showed four bands with a slightly different pattern when compared to the control. The SSCP results of the rest of the *APC* gene exons for all four patients' samples however showed no banding variation compared to the unaffected normal control, suggesting that these other exons do not carry any nucleotide change.

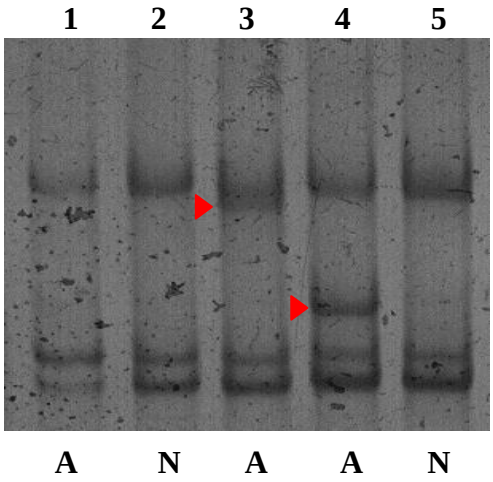

Figure 4.2: SSCP banding pattern of exon 8. Band shifts were observed in lanes 3 and 4 (arrowheads) when compared to the bands in lane 5.

Lane 1: 06-1 Lane 4: 0X-2
Lane 2: 06-2 Lane 5: Unaffected control DNA
Lane 3: 0X-1
(A: samples from individuals clinically diagnosed with FAP; N: unaffected control sample)

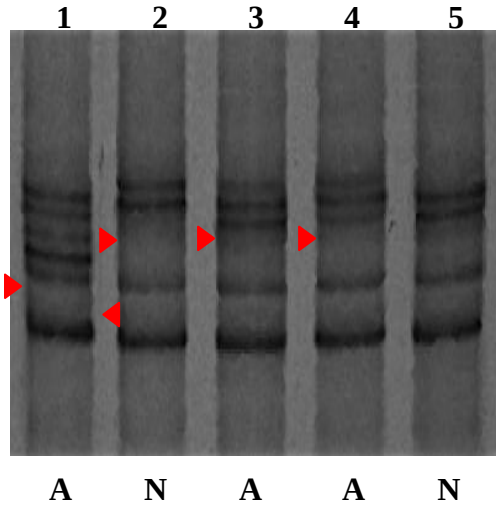

Figure 4.3: SSCP banding pattern of exon 13. Band shifts were detected in lanes 1-4 (arrowheads) when compared to the bands in lane 5.

Lane 1: 06-1 Lane 4: 0X-2
Lane 2: 06-2 Lane 5: Unaffected control DNA
Lane 3: 0X-1
(A: samples from individuals clinically diagnosed with FAP; N: unaffected control sample)

Subsequent SSCP screening on samples 06-3 to 06-7 was carried out to find out whether they carry any nucleotide variation in exon 13. Based on the SSCP banding patterns, samples 06-3, 06-4 and 06-5 exhibited banding patterns similar to that of 06-2, while samples 06-6 and 06-7 showed banding patterns that are unique when compared to that of unaffected control (Figure 4.4).

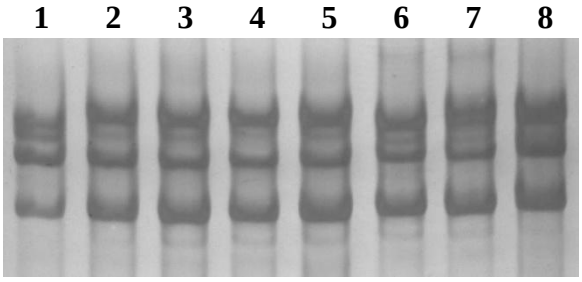

Figure 4.4: SSCP banding pattern of exon 13 of family FAP-06.

Lane 1: 06-1 Lane 5: 06-5
Lane 2: 06-2 Lane 6: 06-6
Lane 3: 06-3 Lane 7: 06-7
Lane 4: 06-4 Lane 8: Unaffected control DNA

4.2: DNA sequencing

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

Sequencing of exon 8 revealed a C>T transition at nucleotide position 847 of the *APC* gene (c.847C>T) (Figure 4.5). This nucleotide substitution creates a new termination codon (CGA>TGA) at codon position 283 in place of the codon that codes for Arginine (p.R283X). The generation of a premature stop codon in the middle of a coding frame is expected to cause an abrupt termination of protein translation process that will produce a truncated APC protein.

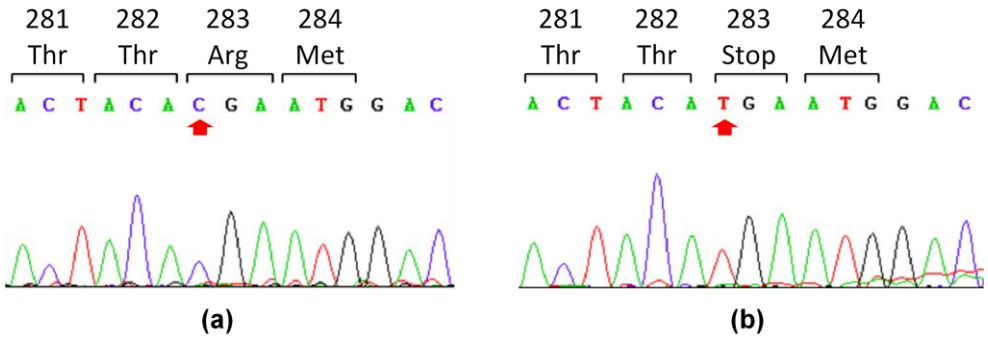

Figure 4.5: Partial chromatogram of cloned exon 8 sequence showing C>T nucleotide substitution (red arrows). **(a)** Partial exon 8 sequence from normal control showing cytosine (C) residue at position 847. **(b)** The corresponding partial exon 8 sequence from sample 0X-1 showing a thymine (T) residue at the same position. Nucleotide substitution at this position creates a stop codon (TGA) at codon position 283.

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

The sequence for exon 8 of sample 0X-2 was identified to carry a single nucleotide insertion between nucleotide position 875 and 876 (c.875_876InsT) (Figures 4.6 & 4.7). The insertion of an extra thymine (T) residue after a stretch of four T's caused the reading frame to shift one nucleotide downstream starting from codon 292 (TTG>TTT). The frame shift will result in a new translation termination signal (TAG) three codons downstream

from the insertion site (p.L292fsX295). The generation of new termination codon in a coding frame is again expected to cause the protein translation process to stop midway and result in truncated APC protein.

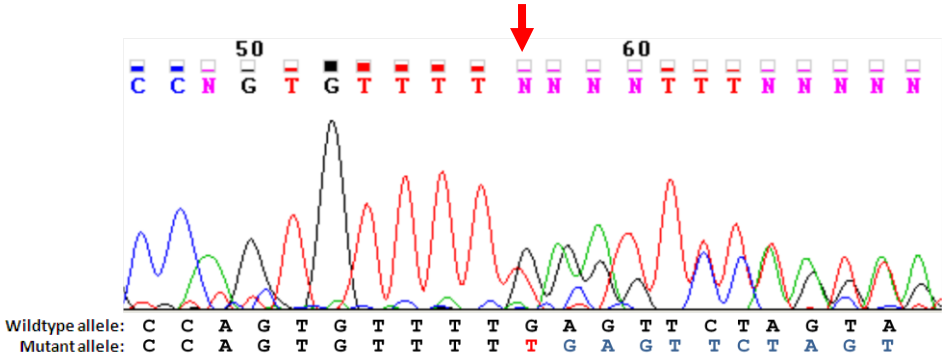

Figure 4.6: Partial chromatogram of PCR product from exon 8 of 0X-2 showing a single thymine insertion (red arrow) that resulted in double peaks from that point onwards.

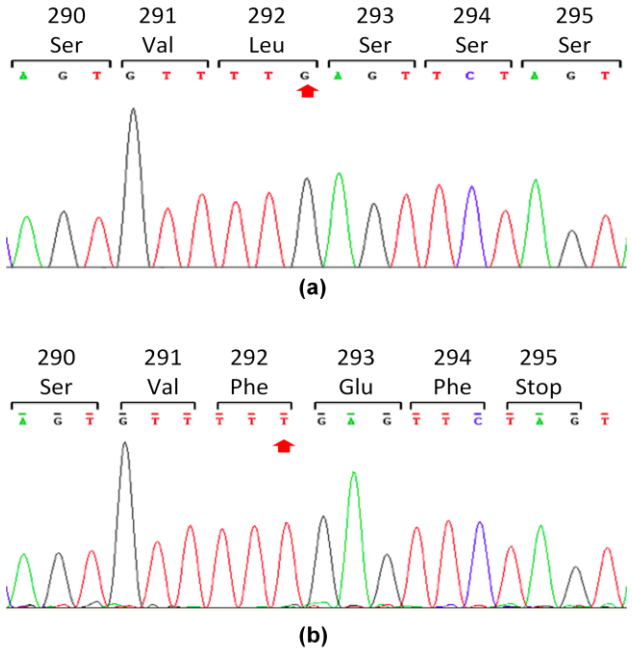

Figure 4.7: Partial chromatogram of cloned exon 8 normal and mutant alleles. **(a)** The sequence of the normal allele showing four thymine residues (spanning codons 291 and 292). **(b)** The sequence of the mutant allele showing an extra thymine (T) residue. This insertion causes frame shift that creates a stop codon (TAG) three codons downstream of the insertion site (codon position 295).

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

Sequencing of exon 13 of sample 06-1 revealed a nucleotide substitution at nucleotide position 1690 of the *APC* gene (Figure 4.8). The C>T transition mutation (c.1690C>T) generate a new stop codon (CGA>TGA) at codon position 564 that replaces the original codon for Arginine (p.R564X). This mutation is expected to cause truncation of the APC protein.

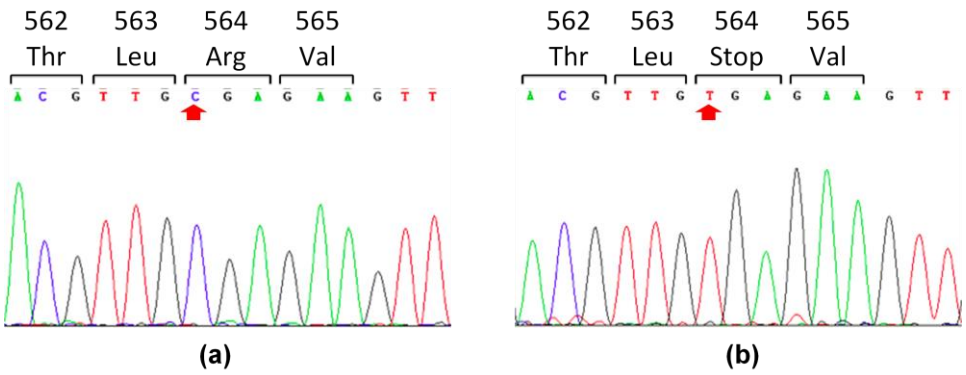

Figure 4.8: Partial chromatogram of cloned exon 13 sequence showing C>T nucleotide substitution (red arrows). **(a)** Partial exon 13 sequence from normal control showing cytosine (C) residue at position 1690. **(b)** The corresponding partial exon 13 sequence from sample 06-1 showing a thymine (T) residue at the same position. Nucleotide substitution at this position creates a stop codon (TGA) at codon position 564.

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

Sequencing analysis of exon 13 for sample 06-7 showed a heterozygous G to A single nucleotide variation at position 1635 (c.1635G>A). Samples 06-2 to 06-5 was found to be homozygous for the G allele while sample 06-6 was homozygous for the A allele (Figure 4.9). The sequencing results were consistent with the banding pattern similarity observed between the samples in the SSCP result (Figure 4.4). The nucleotide variation was unlikely to cause any change at the protein level as both variants would code for amino acid Alanine.

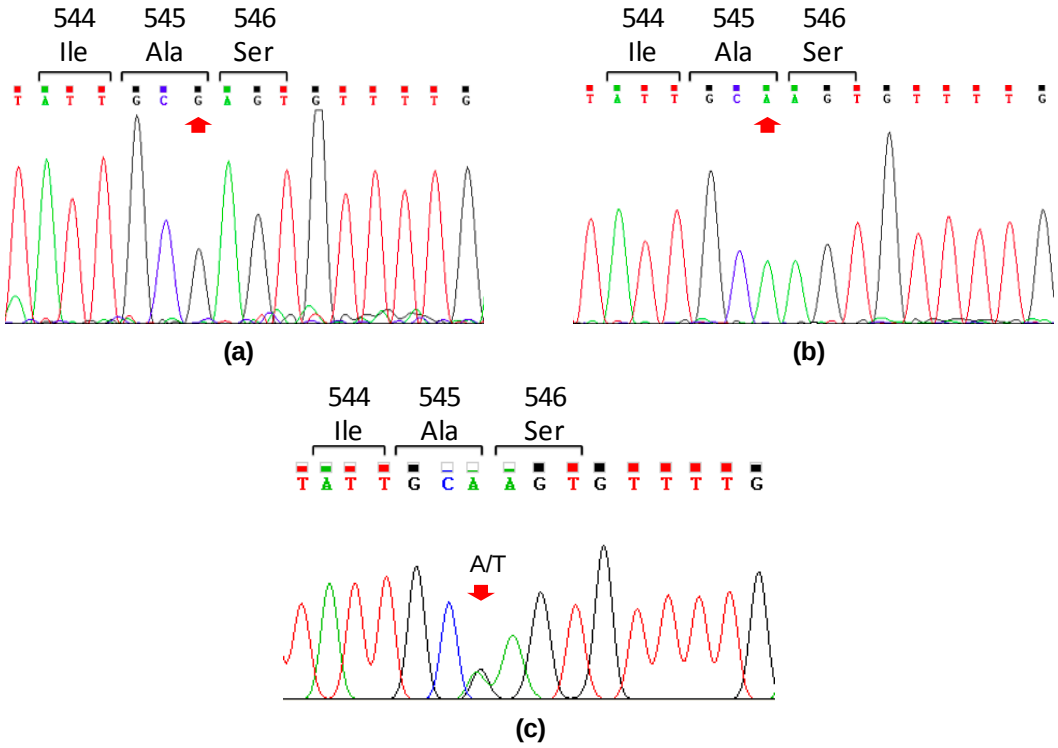

Figure 4.9: Partial chromatogram of exon 13 sequence showing G>A single nucleotide variation (red arrows). Sequences show **(a)** homozygous G, **(b)** homozygous A and **(c)** heterozygous A/G at codon position 545.

4.3: Splicing assay

Splicing assay to assess the effect of R283X mutation on exon 8 splicing was carried out as described in section 3.9. The following subsections address the results obtained from the procedures undertaken.

4.3.1: Minigene construction

Exon 7, exon 8, exon 9a and 9b of the *APC* gene were individually amplified by PCR from DNA sample of an unaffected control. PCR amplification produced amplicons of 440bp, 424bp 692bp and 330bp in size, respectively (Figure 4.10)

Overlap extension PCR was carried out by first overlapping exon 7 with exon 8. This overlap process yielded fragment 7-8 that was 848bp in size (Figure 4.11a). Then fragment 7-8 was joined to exon 9a to produce a 1556bp long fragment (Figure 4.11b). This was followed by fragment 9b splicing to the rest of the construct to produce the 1844bp long construct 7-8-9/7-8m-9 (Figure 4.11c).

4.3.2: Site-directed mutagenesis

Amplification of the insert by PCR after the mutagenesis steps using exon 8 specific primers produced a DNA band that was 424bp in length (Figure 4.12) confirming that the plasmid was carrying the insert of interest. The mutated exon 8 insert was sequenced to confirm that mutagenesis has taken place at the intended nucleotide position. The sequencing result showed that exon 8 was successfully mutated at nucleotidxe position 847 and was free from any spurious mutation (Figure 4.13).

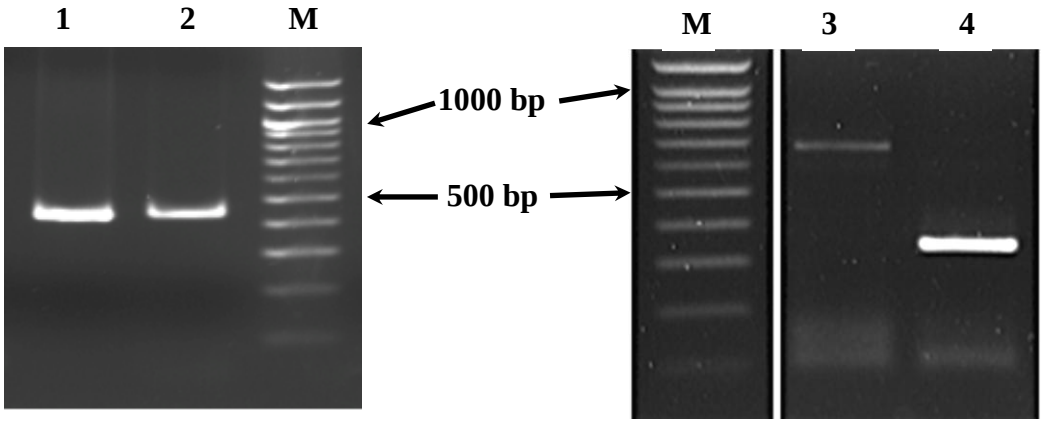

Figure 4.10: Agarose gel electrophoresis of exon 7 (440bp), exon 8 (424bp) and exon 9a (692bp) and exon 9b (330bp) after PCR.
Lane 1: Exon 8
Lane 2: Exon 7
Lane 3: Exon 9a
Lane 4: Exon 9b
Lane M: Marker (100bp ladder)

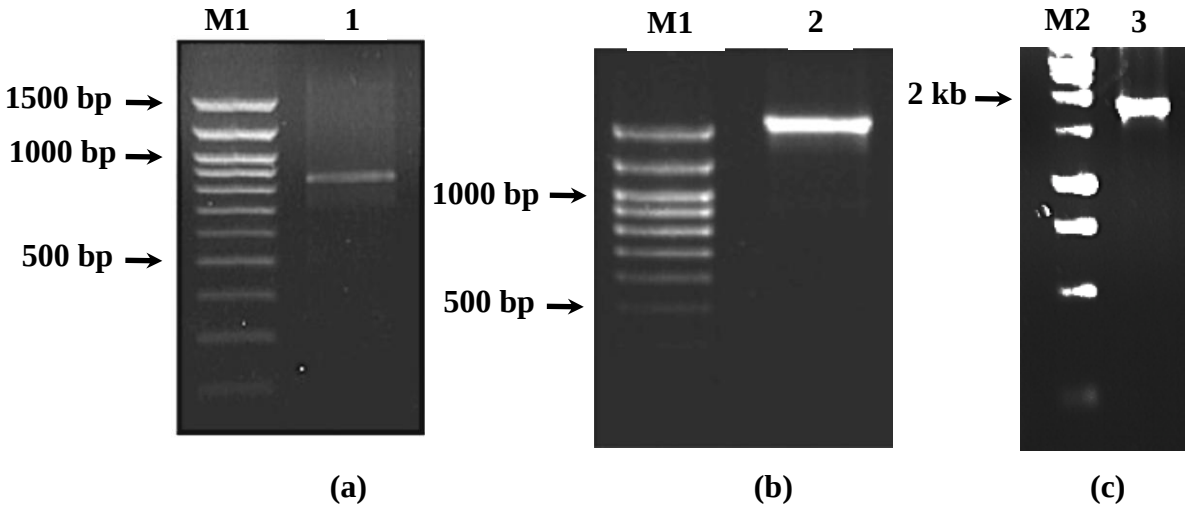

Figure 4.11: Agarose gel electrophoresis of SOE-PCR products. SOE-PCR was carried out to join **(a)** exon 7 and exon 8 to generate 848bp long amplicon, **(b)** fragment 7-8 with 9a of 1556bp in size; **(c)** fragment 7-8-9A with 9b of 1844bp in size.

Lane 1: fragment 7-8
Lane 2: fragment 7-8-9a
Lane 3: fragment 7-8-9a-9b
Lane M1: Marker (100bp ladder)
Lane M2: Marker (1 kb ladder)

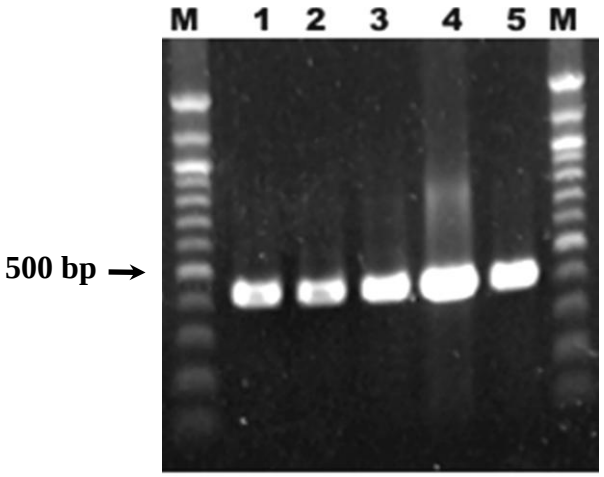

Figure 4.12: Agarose gel electrophoresis of colony PCR after site-directed mutagenesis. The exon 8 inserts were amplified using intronic PCR primers with the expected DNA band size of 424bp.

Lanes 1-5 : Isolated colonies
Lane M: Marker (100bp ladder)

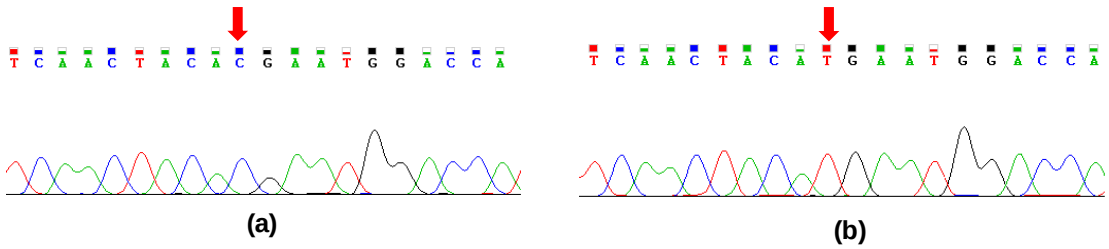

Figure 4.13: Site-directed mutagenesis of exon 8. **(a)** wildtype sequence; **(b)** mutated sequence. C>T mutation is indicated by red arrows.

4.3.3: Restriction enzyme assay

The digestion of pTARGET vector plasmids using *Eco*RI restriction enzyme produced two fragments. The first band with the size of 3kb corresponded to the size of the vector without the insert. The smaller band with the size of 1844bp corresponded to the length of the complete construct which confirms the presence of the minigene construct inside the vector. Figure 4.14 shows the results of plasmid purification and restriction endonuclease digestion assay.

4.4: RT-PCR

RT-PCR was carried out on mRNA samples extracted from the transfected cultured cells. In theory, the transcription product of the mutant minigene is expected to be devoid of exon 8 if the mutation present in exon 8 of the construct does induce exon skipping. Skipping of exon 8 would reduce the length of the expected transcript from 603bp to 504bp.

In Figure 4.15, the RT-PCR result instead showed multiple DNA bands. The expected normal band size was identified as the band approximately 600bp in size. However, another band of similar intensity was detected at around 750bp marker band. The remaining lower intensity bands are also observed at various approximate sizes of 450bp, 1200bp and 1500bp. The results were consistent in all RT-PCR replicates (n=3).

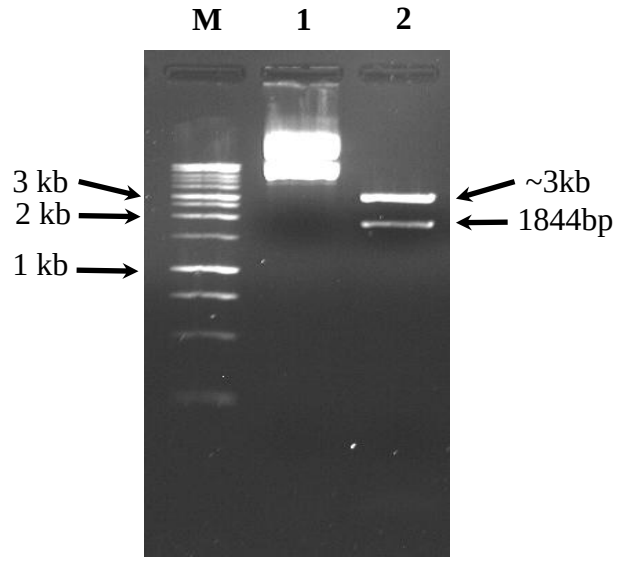

Figure 4.14: Agarose gel electrophoresis of restriction endonuclease digestion assay for insert verification. Purified plasmids were digested using *Eco*RI and produced two bands (3kb and 1844bp).
Lane 1: undigested plasmid
Lane 2: digested plasmid
Lane M: Marker (kb ladder)

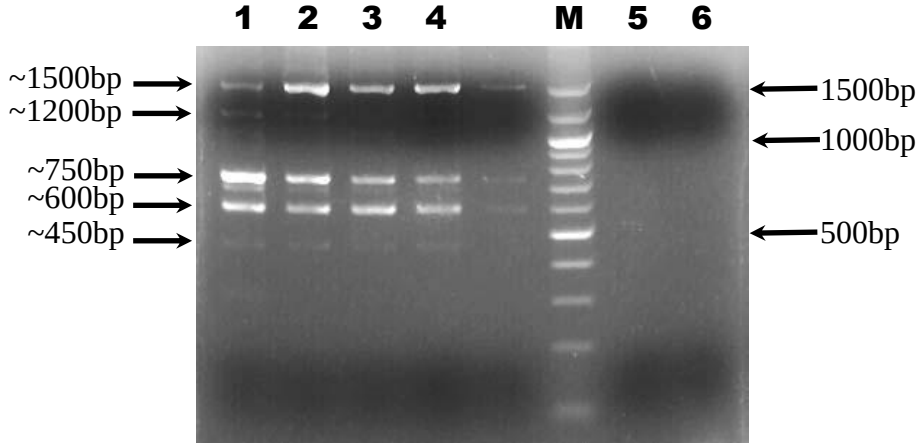

Figure 4.15: Agarose gel electrophoresis of reverse transcriptase (RT-PCR) product of normal and mutant exon 8 constructs expressed in cultured mammalian cells. Both construct produced multiple cDNA amplicons of similar sizes. Approximated product sizes ranges from 450bp to 1500bp in length.
Lane 1&2: cDNA from expressed normal exon 8 construct
Lane 3&4: cDNA from expressed mutant exon 8 construct
Lane 5: negative control for genomic DNA
Lane 6: negative control (minus RNA template)
Lane M: Marker (100bp ladder)

4.5: DNA sequencing and analysis

Sequencing results revealed that all the RT-PCR amplicons described in section 4.4 were in fact products of the *in vitro* expression of the minigene. Figure 4.16 summarizes the structural variation between the splice variants observed. The DNA band corresponding to the molecular weight of ~600bp was identified as the mature mRNA transcript based on its intronless architecture and precise splicing of the exons (Figure 4.17a).

The variation in mRNA lengths however was presumed to be the alternatively spliced mRNA of the minigene. The splicing of these variants followed non-canonical splicing junctions hence the difference in the amplicon length. For the ~460bp DNA band, the transcript was found to utilize an alternative 3' splice site located within exon 9 resulting in the loss of a portion of its 5' end (Figure 4.17b). The ~660bp DNA band was found to utilize both alternative 5' and 3' splice sites located within exon 7 and exon 8, respectively, resulting in the loss of a portion of 3' end of exon 7 and 5' end of exon 8 (Figure 4.17c). The ~750bp band was found to retain intron 6 in the mature transcript (Figure 4.17d). Meanwhile, the ~1200bp DNA band was found to retain introns 6 and 8, hence the increase in length (Figure 4.17e). The sequencing of the ~1500bp band failed to generate result but the band size was consistent with the size of pre-mRNA transcript of the minigene.

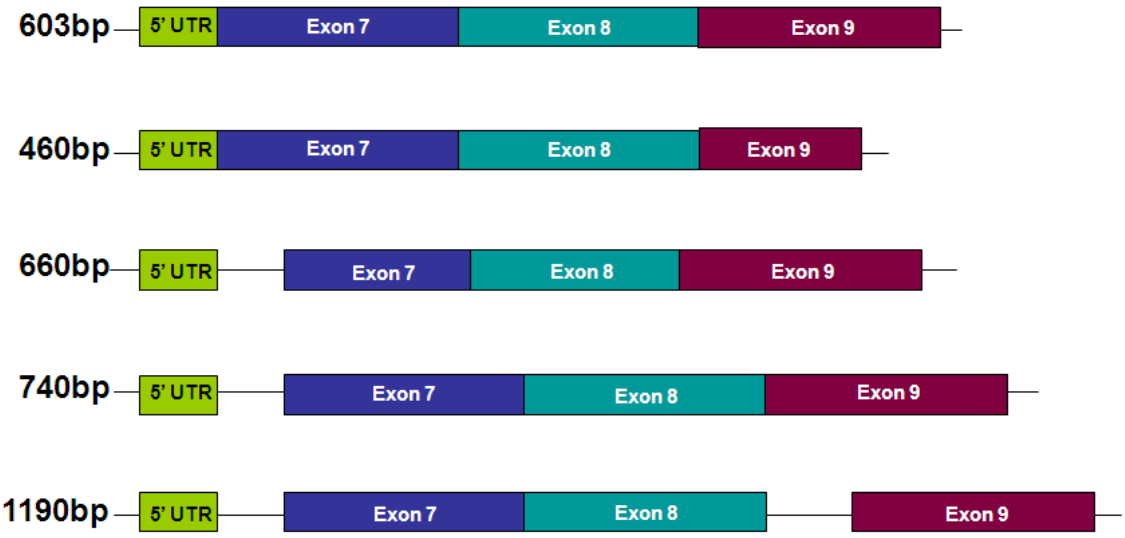

Figure 4.16: Graphical representation of variation in the structure of the mRNA expressed from the minigene. Alternative splicing of the exons were observed with the evidence of intron retention and alternative 3' splice site selection on certain exons.

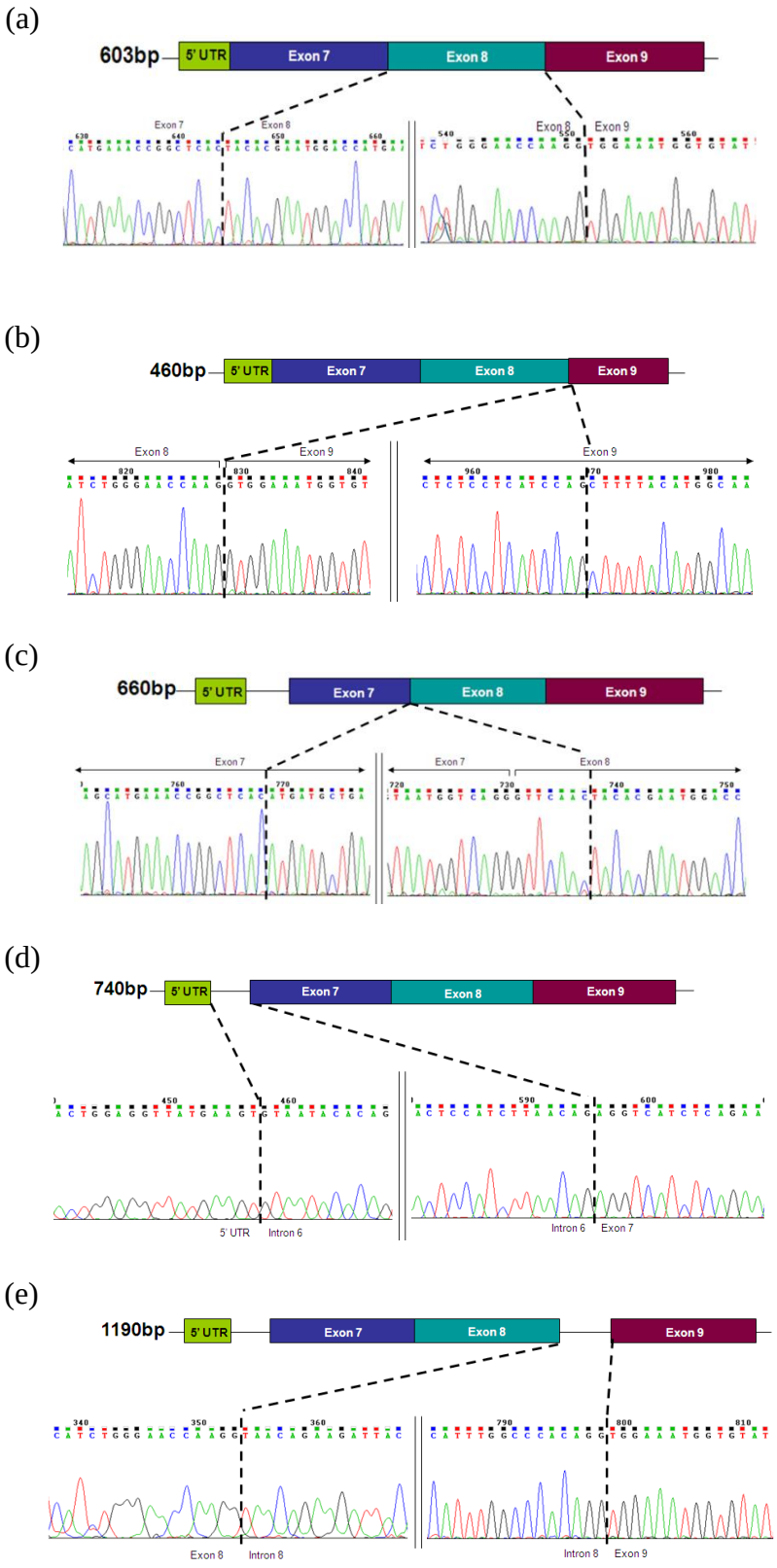

Figure 4.17: Partial chromatograms of the variable splicing products.

5.0: Discussion

Approximately 80% of all CRC cases occur sporadically (Lynch & de la Chapelle, 2003). The remaining 20% are hereditary in nature, and can be dividied into hereditary non-polyposis colorectal cancer (HNPCC), and to a lesser extent familial adenomatous polyposis (FAP) (Haggar *et al.*, 2009). In fact, FAP has the prevalence of less than 1% in total CRC patients compared to 5% for HNPCC (Aretz *et al.*, 2011). The rarity of this condition is reflected by the scarcity of FAP patients available for study. Only three FAP patients were successfully recruited and were used in this *APC* mutation screening study over a period of two years. In addition to low prevalence, the lack of interest of patients (and family members) in participating in research on a disorder that is not immediately treatable or directly beneficial to the patient is also another factor that posed challenges in patient recruitment. Nonetheless, it is generally noted that a larger and ethnically-diverse sampling would be more ideal for such mutation analysis work, especially for epidemiological studies and if any statistical inference such as prevalence and penetrance of the mutations were to be made.

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

In this study, mutation screening of the *APC* gene was carried out by PCR-SSCP analysis. This method had allowed this study to successfully identify the *APC* gene regions that carry the gene mutations in all three of the FAP patients analysed. A total of 38 *APC* gene-coding fragments of the *APC* gene were amplified by PCR from each of the patient's genomic DNA. PCR was carried out using gene-specific PCR primers published in Groden *et al.* (1991) and gave good quality PCR products for SSCP analysis.

SSCP was done using MDE gel that has polyacrylamide-like matrix that is sensitive in detecting DNA conformational differences. According to the protocol supplied by the manufacturer, MDE gel can increase detection of sequence difference up to 80% with faster procedure time compared to polyacrylamide gels. MDE gel concentration of 0.5X was sufficient to screen for mobility shift for most samples. Only in certain cases when unclear mobility shift was observed will the gel concentration be increased by 50%. The high sensitivity of the gel at lower concentration would allow the use of minimal volume of MDE per gel analysis which would translate into cost-saving.

In general, SSCP has been a simple and reliable method for mutation screening for this study. However some problems were encountered during the initial stage regarding the resolution of the DNA bands. The bands at times appeared fuzzy or in smiling/frowning form. Fuzzy bands appearance could be due to many factors including the use of old ammonium persulfate during gel preparation, overheated or incomplete denaturation of samples and the use of NaOH-free loading buffer (MDE Gel Solution Protocol For SSCP and Heteroduplex analyses guide, 2009). These problems however were rectified by using freshly prepared ammonium persulfate during gel preparation as recommended by the troubleshooting guide provided by the supplier.

Apart from that, fluctuation of the gel temperature due to heat generated by the electrical current running through the gel would affect the temperature-dependent conformation of the migrating DNA fragments. This in turn would cause oscillation of DNA conformation during electrophoresis and altogether altering the relative migration rate between samples. Proper dissipation of heat by thermostatically controlled circulator during SSCP is therefore crucial as it would help to maintain the overall optimal SSCP running temperature and prevent the occurence of diffuse or smiling/frowning bands (Hongyo *et al.,* 1993).

Gel staining by silver impregnation to visualize the DNA bands was applied in this study. Silver staining sensitivity is comparable to other non-radioactive methods of visualization (EtBr stains, fluorescent dyes and labels). However, the staining process carried out in this study took approximately 70 minutes to be completed. Recent improvement in silver staining method enabled processing time to be reduced to 20 minutes by eliminating steps involving ethanol, acetic acid and nitric acid treatment before silver staining (Ji *et al.*, 2007). This improved method also allows the use of minimal silver nitrate and DNA concentration (as low as 0.44ng) without compromising staining sensitivity. Such improvements had proved to reduce chemical consumption and helped save time and costs.

Hayashi and Yandell (1993) pointed out that 'one size fits all' concept is not applicable when it comes to considering which SSCP settings to be used. This is due to the multiple factors such as electrophoresis temperature, buffer concentration and additives that according to Suzuki *et al.* (1990) would influence the folding of the single-stranded DNAs. Thus SSCP in general requires optimization for every different DNA fragment to be analysed in order to achieve the highest sensitivity of detection. However this would be a hassle especially when screening need to be done on multiple fragments from many individuals as evident in this study. In this work, eventhough the settings for SSCP electrophoresis were standardized for all DNA fragments, the mutation detection sensitivity in all three of our FAP patients was 100%. This indicates that standardized parameters could be applied in mutation screening SSCP assay. Nonetheless, sensitivity of mutation detection by SSCP is sample-dependant and still requires small trial and error experimentation.

5.2: *APC* mutations

According to the Catalog of Somatic Mutation in Cancer (COSMIC) database (Bamford *et al.*, 2004), frameshift mutations due to insertions (20%) and deletions (41%) form the majority of mutations found in the *APC* gene. This is followed by nonsense mutation (29%) and missense mutations (6.4%).

In this study, three germline mutations in the *APC* gene were identified using PCR-SSCP-direct sequencing method. The first mutation, c.1690C>T was found in exon 13 of patient 06-1 from family FAP-06. This mutation was located in codon position 564 and created an in-frame premature termination codon (p.R564X). The second mutation, c.847C>T was a nonsense mutation found in exon 8 of patient 0X-1. This mutation created a premature stop codon at codon position 283 (p.R283X). Patient 0X-2, on the other hand, was identified to carry a single nucleotide insertion (c.875_876InsT) in exon 8 at codon position 292. This insertion caused the reading frame to shift and created an in-frame stop codon three codons downstream (p.L292fsX295).

The formation of premature stop codons (PTCs) by mutations c.847C>T and c.1690C>T is consistent with the notion that nonsense mutations are the most common germline mutation characterized in the *APC* gene (Rustgi, 2007). Both c.847C>T and c.1690C>T mutations found in this study have been detected in many *APC* gene mutation screening studies. For example, Wallis *et al*., (1999) in their study of *APC* mutation analysis in 190 FAP and 15 non-FAP cases found that c.847C>T mutation is the second most common point mutation observed after the 5bp-deletions at codon 1061 and 1309. Wallis *et al*. (1999) also identified the c.1690C>T mutation found in this study in one of their cases.

A search in the Catalog of Somatic Mutation in Cancer (COSMIC) database revealed that there are six c.847C>T and three c.1690C>T reported cases. The Universal

Mutation Database (UMD) - APC mutation database, on the other hand, reported that c.847C>T mutations have been found in 22 different cases while c.1690C>T mutations were found in 14 cases (Beroud *et al.*, 2000). Similar searches in the LOVD database by the Human Variome Project of the Zhejiang University Center for Genetic and Genomic Medicine (Fokkema *et al.*, 2005) yielded two reports on c.847C>T mutation and 20 cases reported for c.1690C>T mutation.

Further analysis showed that the c.875_876InsT and c.847C>T mutations that occured in exon 8 resided in the DNA stretch that codes for an APC protein region of no known function. This region is located between the heptad repeats within the oligomerisation domain and the Armadillo repeats in the 5' end of the APC protein. Because these truncating mutations are located close to the 5' of the APC protein, the product is expected to lose about 90% of the carboxy-terminal of the protein. This includes losing most of the functional domains including the Armadillo (ARM) repeats, 15 amino acid repeats (15AAR) and 20 amino acid repeats (20AAR), SAMP repeats, domestic domain as well as EB1 and HDLG binding sites.

c.1690C>T mutation however was located slightly downstream of the first two mutations within the codons that code for the ARM repeats. This nonsense mutation would cause the protein synthesis to terminate in the middle of this repeat region. However the protein produced would still lose the majority of the downstream functional domains. The loss of these domains would collectively disable the negative regulation of the Wnt signalling pathway (Morin *et al.*, 1997). Ultimately, it would cause hyperactivation in the transcription of Wnt target genes which is deemed as one of the hallmark of colorectal carcinogenesis (Gavert and Ben-Ze'ev, 2007). Furthermore, the partial retention of the ARM in the case of c.1690C>T might be favored in tumorigenesis. Truncated APC protein with retained partial ARM region has been

shown to contribute to Asef-mediated aberrant tumour cell migratory properties (Akiyama and Kawasaki, 2006).

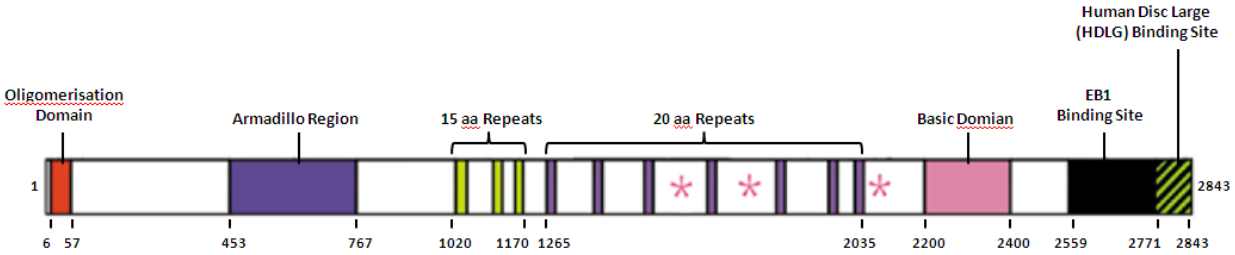

Figure 5.1: APC protein structure. *(Source: Fearnhead* et al.*, 2001)*

In accordance with the Knudson 'two hits hypothesis', cells with germline mutations in the *APC* gene would accumulate their second 'hits' over time and result in loss of heterozygous state and function of both alleles. However, earlier assumption that both mutational events occur independently was challenged with evidence of positional interdependence between germline and somatic mutations (Shoemaker *et al.*, 1997). Albuquerque *et al.* (2002) reported that the type and position of inactivating somatic mutations in the *APC* alleles of colonic tumour cells are biasly selected towards ensuring only one of the *APC* allele is left with one to two intact 20AARs. They hypothesized that by having only one to two 20AARs, the transforming colonic epithelial cells could retain β-catenin degradation activity at a certain level that favors tumorigenesis as oppose to a complete inactivation of β-catenin degradation. Similar first hit-second hit relationship was also observed in desmoid tumours, an extracolonic feature of FAP (Latchford *et al.*, 2007).

The heterozygous state of the mutations detected in this study was further supported by the analysis of the sequence chromatogram of the mutated exons that showed the presence of sequence with normal genotype. Based on this, carcinogenic somatic mutations in all the FAP patients in this study are conjectured to have occurred in the mutation cluster region (MCR) of the *APC* gene. However, the unavailability of

tumour samples from these patients had precluded any effort to corroborate the inference made.

A similar analysis was also carried out for the third germline mutation found in this study (c.875_876InsT). Unfortunately however, searches in various databases failed to identify any matches to this mutation suggesting that this mutation could be a novel *APC* mutation (see Appendix D). In order to confirm the pathogenicity of this mutation, ideally one would perform experiments to elucidate the functional properties of the mutant protein. Analysis such as *in vitro* protein synthesis of *APC* mRNA carrying this mutation as well as protein-protein binding assay such as the yeast two-hybrid experiments could also be carried out to verify the effect of this mutation on the structure and function of the APC protein respectively. Alternatively, the involvement of this mutation in FAP can be inferred based on its absence in healthy unaffected individuals. This work would involve testing a set of sample representing the general population for the presence of this mutation. Due to time constraints, this work could not be carried out in this study, but would indeed be a necessary component for future studies before definitive conclusion on its causal role can be made.

5.3: The genotype-phenotype correlation of extracolonical FAP features

As reviewed by Nieuwenhuis and Vasen (2007), there have been extensive reports demonstrating how FAP phenotypic outcome can be linked to its *APC* genotype. The severity of FAP often times correlates with the position of the mutation in the *APC* gene especially in the MCR region that would result in profuse polyp formation. Mutations that occur in the immediate proximal and distal region of the MCR however are implicated with an intermediate phenotype while mutations in the far 5' and 3' ends of the gene are responsible for the attenuated form of FAP. The mutations identified in this study were located in the proximal region of the MCR suggesting that the carriers would

be affected with intermediate classical polyposis, which is consistent in their clinical diagnosis report.

In addition, patient 06-1 was reported to have periampullary adenomatosis, a form of extracolonic feature of FAP. It has been shown that colorectal cancer patients would have an elevated risk to develop a second primary ampullary malignancy (Das *et al.,* 2004). Esposito *et al.* (2001) concluded that genetic alteration in *APC* and K-*ras* genes probably form the initial pathogenic step by triggering adenomatous growth and malignancy in the ampullary region similar to colorectal tumours. In a different study, Kim *et al.* (2007) demonstrated that hypermethylation of the *APC* and *E-cadherin (CDH1)* genes also contributed to carcinogenesis in the ampullary duct.

There have been contradicting data regarding the relationship of specific *APC* mutations with the occurrence of periampullary adenocarcinoma. While some claimed that specific germline mutations in FAP families have no correlation with periampullary adenoma development despite showing familial segregation of the frequency and severity of the malignancy in FAP families (Sanabria *et al.,* 1996; Kashiwagi and Spigelman, 2000), there are strong evidences linking periampullary carcinogenesis to mutations in codons 1000 to 1700 (Attard *et al.,* 2004; Bjork *et al.,* 2001; Andersen *et al.,* 1999; Leggett *et al.,* 1997; Toyooka *et al.,* 1995) which also coincided with the location of the MCR in *APC* exon 15. Based on this, we postulated that patient 06-1 may have acquired a somatic mutation in the MCR region that, in synergy with the c.1690C>T germline mutation found from this screening, would result in development of the periampullary adenomas. The segregation of this phenotype or the mutation within the family FAP-06 however cannot be confirmed due to lack of clinical data. Extraintestinal manifestation of FAP however was not observed in 0X-1 and 0X-2.

5.4: c.1635G>A polymorphism

A single nucleotide variation, c.1635G>A was found in exon 13 of the healthy control sample. It is located nine nucleotides downstream of the 3' splice site of exon 13. The G>A transition was found to be synonymous for Alanine codon. This SNP has been reported in different populations (Zhou *et al.*, 2004; Hadjisavvas *et al.*, 2006) while there is no record of this SNP in the Malaysian population as yet.

It is interesting to note that a report by Venesio *et al.* (2007) have used the A allele as the wildtype allele when describing this particular polymorphism. The author later proposed that the predicted outcome of this "polymorphism" could be the disruption of an SRp40 splicing motif and cause exon skipping. The fact that the author identified the G as the mutated allele conflicted with many SNP databases that use G allele as the wildtype allele (for example reference sequence accession number: NG_008481.4). While the discrepancy on the wildtype status of an allele can be easily resolved by screening adequate numbers of normal controls in a particular population, such claims of the potential pathogenecity of silent mutations (in this case, its effect on splicing) obtained through *in silico* studies need to be corroborated by an appropriately designed splicing assay and mRNA analysis.

5.5: *In vitro* splicing assay

In 1998, Laurent-Puig and colleagues estimated that around 98% of *APC* mutations lead to protein truncation either by frameshift (62%) or nonsense mutations (34%), while the effect of missense mutations are often related to the protein function. Silent mutations on the other hand are just considered as benign polymorphisms. However with the discovery of signalling regions in the pre-mRNA transcripts that regulate their splicing outcome, non-truncating mutations are increasingly being

implicated with pathogenecity of disease, mainly due to causing aberrant splicing of mRNA (reviewed in Caceras and Kornblihtt, 2002; Cartegni *et al.*, 2002).

Mutation in exonic splicing enhancer (ESE) sites in particular has been found to be pathogenic in many disorders. Apart from the weakly conserved 5' and 3' splice sites, ESE motifs are signalling components that are important in distinguishing exon-intron boundaries in mRNA splicing. Abrogated ESEs have been shown to cause skipping of exons, which would commonly cause frame shifts that may generate in-frame premature stop codons (PTC), as demonstrated in an FAP case described by Goncalves *et al.* (2009).

However in a case of PTC-bearing ESE in exon 8 of the *APC* gene, skipping of exon 8 due to its ESE disruption might prove to be an advantage. Theoretically, skipping of exon 8 due to splicing error or by the putative action of nonsense-associated altered splicing (NAS) would not change the original reading frame and is predicted to allow the synthesis of an 'almost complete' APC protein. Exon 8 of the *APC* gene codes for a protein portion in the 5' region of the APC protein that is not part of the main functional domains in the protein. Therefore the portion missing due to skipping of exon 8 would be considered as non-detrimental to protein function. Open reading frame (ORF) -preserving exon skipping has been observed, among others, in mRNA analysis of dystrophin gene mutations that cause Becker muscular dystrophy (BMD), a milder variant of Duchenne muscular dystrophy (DMD) (Shiga *et al.*,1997).

The effect of the ESE-disrupting mutation in exon 8 of the *APC* gene on constitutional splicing was tested by using a minigene system. *APC* minigene was constructed and expressed in HepG2 cultured cells and analysed by RT-PCR. Since *APC* gene is endogenously expressed in the host cells, RT-PCR was done using forward primer of the vector and gene specific reverse primer. This is to ensure that the PCR was amplifying specifically from the *APC* mRNAs expressed from the expression

vector. It was expected that this exon would be skipped during post-transcriptional RNA processing due to the nullified ESE motif of exon 8 and produce a shorter isoform (503bp) of the mature *APC* mRNA. But the result of cDNA analysis however did not show the spliced product expected.

The absence of the shorter *APC* mRNA variant lacking exon 8 indicated that the mutation might not influence alternative splicing as expected despite the ESE disruption predicted *in silico*. It has been shown that one of the function of ESEs is to support exon boundary definition at a weak 3' splice acceptor site and that an engineered mutation to strengthen the splice site succeeded in diminishing the necessity of the adjacent ESE function (Fukao *et al.*, 2010). ESE function, therefore in this case, may not be critical for proper exon 8 splicing due to its strong 3' splice site consensus sequence.

Conversely, alternative mRNA splicing of exon 8 might demonstrate specificity towards cell type. In this study, aside from cDNA with retained exon 8, the presence of cDNA variants with intronic retentions proximal to exon 7 and exon 9 as well as the use of alternative splice site for exon 9 was also observed. This outcome was suspected to be the effect of tissue-specific modulation by variable concentration of enhancing and antagonizing splicing factors (Caceras *et al.*, 1994). Splicing assay of a minigene construct expressed in different cell types has been found to give rise to variable amounts of canonically spliced transcripts (Pagani *et al.*, 2003). The quantitative presence of co-factors that mediates spliceosome complex interaction with ESE may also contribute to the cell-specific effect (Seong *et al.*, 2002). Taking this factor into consideration, the use of colorectal cancer cell lines in the *APC* exon 8 minigene splicing assay instead of hepatocarcinoma cells would probably give a better representation of the *in vivo* mutated-ESE splicing event.

6.0: Conclusion

This mutation screening study of the *APC* gene has successfully found three mutations (c.847C>T, c.1690 C>T and c.875_876InsT) which are potentially pathogenic in Malaysian FAP patients. Mutations c.847C>T and c.1690 C>T have been reported before but no reports were found on mutation c.875_876InsT suggesting that this mutation might be a novel APC mutation. A genotype-phenotype relationship between the loci of these *APC* mutations and its FAP phenotype (polyposis severity and extracolonic tumour occurrence) was consistent with previous reports. The protein-truncating nature of the mutations observed was also consistent with other *APC* mutations observed in other reports. This supports the postulation that these mutations are the causative agent of FAP in this set of FAP patients. With this information, early genetic testing and appropriate genetic counseling sessions for the family members of the patients can be carried out to reduce their risk of developing FAP later in life.

The splicing assay carried out in this study demonstrated how the abolishment of an ESE motif by mutation does not necessarily translate to altered splicing process. Although expected to cause exon skipping by splicing prediction software, the abolishment of the ESE-motif by c.847C>T mutation did not show the expected skipping event of exon 8. This outcome supports the need to corroborate data obtained from splicing prediction software with experimental data from *in vitro* splicing assay.