MODIFIER GENE POLYMORPHISMS AND INFLUENCE ON DISEASE EXPRESSION IN FAP

Thesis for the degree Master of Pharmacy

By

Eva Kristin Andreassen
2009

Institute of Pharmacy
Faculty of Medicine
University of Tromsø

Supervisors:
Professor Rodney J. Scott, School of Biomedical Science, University of Newcastle
Professor Thrina Loennechen, Institute of Pharmacy, University of Tromsø

ACKNOWLEDGEMENTS

This master thesis was carried out at Hunter Medical Research Institute, John Hunter Hospital in Newcastle, Australia from October 2008 to May 2009 under supervision of Professor Rodney J. Scott and Professor Thrina Loennechen due to collaboration between the University of Newcastle and the University of Tromsø.

First of all I would like to thank the University of Newcastle for giving me the opportunity to work with this interesting project. I would like to thank my supervisor Professor Rodney J. Scott for the help and advice while working with this thesis. I owe a special thanks to Dr. Katie Ashton for all the help and guidance she has offered along the way. I would like to thank Dr. Bente Talseth-Palmer for making the arrangements for me to come to the University of Newcastle and for her advice while working with this project. Furthermore I would like to thank the rest of the staff/students at Hunter Medical Research Institute, department of Information based medicine for help and support while I have been at the institute.

A thanks goes to my in-house supervisor Professor Thrina Loennechen for the advice and help she has offered while I have been working with this project.

I special thanks goes to Ingvild Synnøve Brenne for her friendship and support through this period. A good friend is priceless when you travel across the world to learn and experience new things.

Finally I would like to thank my family for their encouragement and support through my years of study. They have given me love, understanding and the confidence to go my own way. I could not have finished this thesis without them.

Newcastle, May 2009 Eva Kristin Andreassen.

ABBREVIATIONS

AA	Homozygous wildtype genotype
Aa	Heterozygous variant genotype
Aa	Homozygous variant genotype
AFAP	Attenuated familial adenomatous polyposis
APC	Adenomatous polyposis coli
ATP5a1	ATP synthase gene
BER	Base excision repair
Cdk4	Cyclin dependent kinase 4
CHRPE	Congenital hypertrophy of the retinal pigment
CK1	Casein kinase 1
CK2	Casein kinase 2
c-Myc	v-myc myelocytomatosis viral oncogene homolog (avian)
CRC	Colorectal cancer
CRAC1	Colorectal adenoma and carcinoma 1
DNA	Deoxyribonucleic acid
EIF3H	Eukaryotic translation initiation factor 3, subunit H
EDTA	Ethylenediaminetetraacetic acid
FAP	Familial adenomatous polyposis
GI	Gastro intestinal
GREM1	Gremlin 1, cysteine knot superfamily, homolog (Xenopus laevis)
GSK3β	Glycogen synthase-3β kinase
G382D	Glycine at position 382 replaced with aspartate
HNPCC	Hereditary non-polyposis colorectal cancer

LEF/TCF	Lymphoid enchancer-binding factor 1/ T cell specific transcription factor
LOC120376;FLJ45803	Hypothetical protein LOC120376;FLJ45803 protein
LOH	Loss of heterozygosity
MAP	MYH-associated polyposis
MCR	Mutation cluster region
Min	Multiple intestinal neoplasia
MMR	Mismatch repair
Mom	Modifiers of Min
MUTYH	MutY human homologue
MYH	MutY human homologue
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
PCR	Polymerase chain reaction
POU5F1P1	POU class 5 homeobox 1 pseudogene 1
RT-PCR	Real time polymerase chain reaction
SMAD7	SMAD family member 7
SNP	Single nucleotide polymorphism
SPSS	Statistical Package for the Social Sciences
Y165C	Tyrosine at position 165 replaced with cysteine

ABSTRACT

Background: Familial adenomatous polyposis (FAP) is a hereditary autosomal dominant condition characterized by the appearance of hundreds to thousands of adenomatous polyps throughout the colon and rectum. If the condition is left untreated it will ultimately develop into colorectal cancer. Most of the articles available today concerning FAP focus on the adenomatous polyposis coli gene as this gene appears to be the most frequent mutated gene in FAP patients. Recent studies suggest that certain modifier genes may play an important role in the development of colorectal cancer in the general population this may also be true for FAP.

Aim of the study: The aim of this study was to examine 199 Australian APC mutation positive FAP patients with a molecular diagnosis of FAP for polymorphisms in five loci on chromosomes 10p14, 8q23.3, 8q24, 11q23, 18q21 and the gene ATP5a1.

Methods: The genotypes were determined for each individual by finding the polymorphic combination for the candidate single nucleotide polymorphisms (SNPs). When a genotype was determined for a particular DNA sample, the nucleotide present at the polymorpic site was reported. The results were analysed using statistical programs.

Results: The result of this study revealed that there is an increased risk of developing polyps if a patient harbours the heterozygote or variant genotype of the SNP rs10318 compared to patients with the wildtype genotype. This appears to be a protective effect against early polyp development in FAP individuals who harbours the variant genotype of SNP rs3802842 compared with those with a wildtype or heterozygote genotype. Patients harbouring the variant genotype of the SNP rs4779584 have an increased risk of developing polyps and CRC at an earlier stage than patients harbouring a wildtype or heterozygote genotype.

Conclusion: In conclusions the findings from this study indicate that modifier genes have an effect on disease expression in FAP patients. These results warrant further investigation in larger FAP populations that harbour APC mutations to confirm the associations reported herein.

TABLE OF CONTENTS

1 - INTRODUCTION

1.1 Colorectal cancer

Colorectal cancer (CRC) is one of the most common and preventable forms of cancer worldwide, accounting for more than 100,000 deaths the United States in 2008 [1]. Several genetic and environmental factors contribute to the development of cancer and it is estimated that up to 35% of all colorectal cancers are caused by a genetic predisposition [2]. During the years of research several high penetrance alleles have been identified [3]. These include alleles causing Familial Adenomatous Polyposis (FAP), Hereditary Non-polyposis Colorectal Cancer (HNPCC/Lynch syndrome), Peutz Jeghers syndrome and Juvenile polyposis. Genetic testing is available for these conditions and there are different protocols for how the patients and their families should be treated [2]. The focus of this thesis is on the condition FAP which accounts for approximately 1% of all colorectal cancers [4].

1.2 Familial Adenomatous Polyposis

Familial adenomatous polyposis (FAP) is a hereditary autosomal-dominant condition characterized by the appearance of hundreds to thousands of adenomatous polyps throughout the colon and rectum [4]. Manifestation outside the colon occurs frequently and include Gardner's syndrome, congenital hypertrophy of the retinal pigment epithelium, hepatoblastoma, fundic gland polyps in the stomach, pancreas and thyroid, dental abnormalities and malignant tumours in the central nervous system [4, 5]. If left untreated the condition will ultimately develop into colorectal cancer [4]. However, the risk of developing cancer is generally considered to be related to polyp number, with more polyps leading to a higher risk of developing cancer [4].

The incidence of FAP in the population is approximately 1 in 8000 and the polyps usually appear by adolescence or third decade of life [5]. Adenocarcinomas occur 10-20

years after the polyps have developed and the majority of the carcinomas are found on the left side of the intestine [5].

Genetic testing is important for FAP patients seeing as the vertical transmission through generations is very likely due to the condition's autosomal dominant inheritance pattern with high penetrance [2]. Patients with the following characteristics should be tested [2]: 1) Individuals with some but not all the features of classical FAP. 2) Individuals who manifest clinically defined FAP, but where the mutation has not been defined within the family (up to 25% of FAP cases come from de novo APC mutations and these patients will have no family history of the disease). 3) Relatives within an FAP cohort family once the founder mutation is known.

When a patient presents with colonic polyps, there is a strong recommendation to have an annual colonoscopy from adolescence considering the high risk of developing colorectal cancer which is often followed by prophylactic colectomy or proctolectomy [4].

1.3 Attenuated Familial Adenomatous polyposis

Attenuated familial adenomatous polyposis (AFAP) is a milder form of FAP where there is a later diagnosis of polyps (>50 years), fewer polyps (5-100), more frequent localization in the proximal colon and fewer extracolonic manifestations [2, 5]. APC germline mutations are found in 20-30% of patients with AFAP and the mutations are often located at the proximal or distal parts in exon 9 of the APC gene [6].

The frequency of upper gastro intestinal (GI) polyposis and other features of AFAP are similar compared to classical FAP [6]. Depending on the age of onset of colorectal cancer in a AFAP family, surveillance may be initiated later than in families with classical FAP [6]. Colonoscopy is usually offered by the age of 15 [6].

AFAP patients carry a significant increased risk of developing colorectal cancer and is sometimes mistaken for sporadic cases of colorectal cancer or hereditary non-polyposis colorectal cancer (HNPCC) [5].

1.4 Extracolonic manifestations

Extracolonic manifestations occur in approximately 70% of FAP patients [7]. It still remains to be determined whether the manifestations can be related to the location of the mutations in the APC gene [7]. The different extracolonic manifestations include Gardner's syndrome, desmoids disease, congenital hypertrophy of the retinal pigment epithelium, hepatoblastoma, fundic gland polyps in the stomach, pancreas and thyroid, dental abnormalities and malignant tumours in the central nervous system [4, 5, 8, 9].Gardner's syndrome is a condition where the colonic polyps are associated with epidermoid skin cysts and benign osteoid tumours of the mandible and long bones and is very common in cases of FAP [5]. Desmoid disease are rare benign and invasive firbromatoses [8]. They appear in 10% of FAP patients and are prone to local invasions and recurrence but tend not to metastasize [8]. Congenital hypertrophy of the retinal pigment epithelium (CHRPE) occur in ~60% of FAP families [4]. The condition can be detected by opthalmoscopy at any age and is therefore a good technique to use to identify at-risk families before the polyps have appeared [4]. CHRPE is a condition without malignant potential which does not affect the patient's sight [4]. FAP patients also have a significantly increased risk of hepatoblastoma, fundic gland polyps in the stomach, pancreas and thyroid, dental abnormalities and malignant tumours in the central nervous system known as Turcot's syndrome [4, 10].

1.5 The genes involved in the development of FAP

The genes known to be involved in the development of FAP include the adenomatous polyposis coli (APC) and the MutY homolog (MYTYH) genes [4]. However, recent studies suggest that certain modifier genes may also play an important role in the

development of CRC [11-13]. Most of the literature concerning FAP focuses on the APC gene as this gene is the most frequent mutated gene in FAP patients. Research show that 16% of the APC-mutation-negative FAP patients have a mutation in the MUTYH gene [14].

Several articles recently published focus on modifier genes and the effect they have on development of CRC [11-13]. Another recent study show that the gene known as ATP5a1 may act as a modifier gene in the development of colorectal cancer [15]. Since these recent findings do not address the condition known as FAP the main focus of this thesis will be on establishing whether the polymorphisms in five loci on chromosomes 10p14, 8q23.3, 8q24, 11q23, 18q21 and the gene ATP5a1 may have an influence on the development of FAP and CRC.

1.5.1 The Adenomatous Polyposis Coli gene

FAP is caused by a germline mutation in the tumour suppressor gene adenomatous polyposis coli (APC) on chromosome 5q21-22 [16]. Most of the germline mutations are nonsense or frameshift mutations and result in a truncated protein product with abnormal functions [4].

According to Knudson's two-hit hypothesis, FAP patients develop somatic APC mutations or loss of heterozygosity (LOH) in addition to the original germline mutation they carry [4]. The site of the "first hit" which is the germline mutation seems to determine the "second hit" which is the somatic mutation in both FAP and sporadic cases of CRC [17].

There is a phenotype-genotype correlation occurring in APC positive FAP patients [4]. Polyp number in FAP patients appears to depend on the location of the mutation in the APC gene [16]. However, this is still uncertain as there are many factors contributing to the development of colorectal cancer [4]. Generally, germline mutations located between

codons 450 and 1600 result in a more stable truncated protein. These mutations seem to be associated with a higher number of polyps [18].

The APC gene consist of 8535 base pair organized into 21 exons [4]. Most of the germline and somatic APC mutations occur in exon 15. More than 50% occur between codons 1286 and 1513 [19]. This region is known as the mutation cluster region (MCR) [19]. Since this region is frequently affected by mutations, it is one of the first areas in the APC gene that are searched in when new families are tested for APC mutations (see figure 1.1) [5]. APC mutations in the MCR region seems to be associated with LOH, while mutations that are not in the MCR, seem to be associated with protein truncating mutations [4].

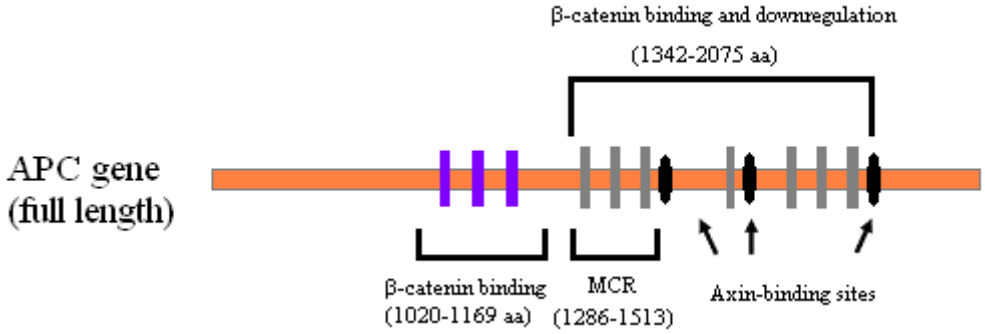


Figure 1.1 Domains present on the APC gene. The mutation cluster region can be seen from codon 1286 to 1513 and is a hot spot for mutations in the APC gene (figure adapted from Narayan and Roy [10]).

As one of the earliest events in the multi-step development of colorectal cancer, the APC gene mutation plays an important role in cell growth and migration, signal transduction, and control of chromosome stability [10]. APC regulates β-catenin levels through Wnt-signaling, thus the Wnt-signaling pathway plays an important role in the development of colorectal cancer [10].

The Wnt signaling pathway (figure 1.2) is important in organ development, cellular proliferation, morphology, motility, and embryogenesis [10]. APC seems to down regulate β-catenin through the Wnt-signaling pathway [10]. β-catenin signals in the transformation of colonic epithelial cells and in melanoma progression and is believed to be the main partner of APC in its tumour suppression function [10, 19].

β-catenin levels are normally low in the cytoplasm through proteasome-mediated degradation [20]. This degradation is controlled by a complex containing axin (Axil/conduction), APC, and glycogen synthase-3β kinase (GSK3β). When cells receive Wnt signals, the GSK3β is inhibited from phosphorylating the β-catenin, APC and Axin complex, thus β-catenin accumulates in the cytoplasm and nucleus [10, 20].
β-catenin in the nucleus will interact with transcription factors such as lymphoid enhancer-binding factor 1/T cell specific transcription factor (LEF/TCF) and affect the transcription [20]. Consequently, if there is a mutation of β-catenin or truncation of APC, the stability and transcriptional activity of β-catenin increases [10]. Due to the increased cell proliferation in the colonic epithelial cells adenomatous lesions will appear [20].

Recent observations suggest that colon tumours carrying mutations in the APC gene also have increased levels of c-Myc, a proto-oncogene important in cellular proliferation [10, 21]. C-Myc seem to primarily function as a transcriptional regulator and is involved in numerous critical processes including apoptosis, proliferation, cell metabolism, DNA repair, and angiogenesis [21].
C-Myc also functions as a transcriptional repressor of genes involved in cell adhesion and growth arrest. Recent data even suggest that c-Myc might play a direct role in DNA replication [21]. When c-Myc is increased through the Wnt-signaling pathway, the expression of cyclin dependent kinase 4 (Cdk4) gene is up-regulated. The Cdk4 gene produces a product responsible for cell cycle regulations in the G_1 phase [10].

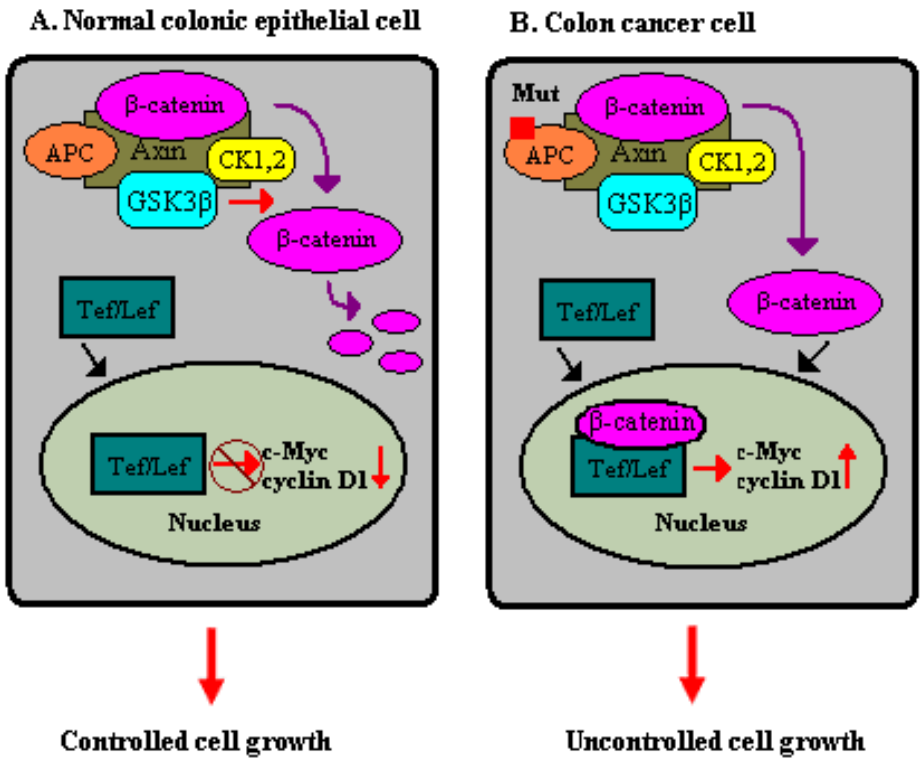


Figure 1.2 A model of the Wnt-signaling pathway. Figure A shows the normal colonic epithelial cell. In the absence of Wnt-signaling, GSK3β and CK1 or 2 kinases become active and phosphorylate β-catenin [10].

Figure B shows the colon cancer cell. There is an increased level of β-catenin due to the mutation in APC gene or β-catenin (figure adapted from Narayan and Roy. [10]).

1.5.2 MutY human homologue gene

Thousands of DNA lesions are repaired every day by different DNA repair systems that have overlapping specificities [22]. At present, approximately 150 DNA repair gene products have been identified [22, 23]. There appears to be 5 major repair pathways: 1) base excision repair (BER), 2) mismatch repair (MMR) 3) nucleotide excision repair (NER) 4) homologous recombination (HR) and 5) non-homologous end joining (NHEJ) [24]. The five repair mechanisms function together to protect DNA from environmental attack and keep normal cellular metabolism to maintain genome integrity [24].

MutY human homologue (MUTYH) gene is a base excision repair gene playing an important role in the process of repairing DNA damage caused by reactive oxygen species [14]. If the MUTYH protein is dysfunctional, mutations in the APC gene may occur, the most frequent mutations being G:C→T:A transversions [25]. The MUTYH gene is located on chromosome 1 between p32.1 and p34.3 [25]. There are two hot spot mutations responsible for 80% of the mutations in the MUTYH gene, Y165C and G382D [6]. Research show that 10-20% of patients with APC-mutation-negative FAP and AFAP harbour a bialleic mutation in the MUTYH gene [6]. This condition is called MYH-associated polyposis (MAP) and can be difficult to distinguish from classical FAP and AFAP because of the similar phenotype [6].) After excluding APC gene mutations in FAP patients by genetic testing, MUTYH germline mutation testing is indicated [6]. Approximately 50% of patients with MUTYH gene mutations develop colorectal cancer [25].

1.6 Single Nucleotide Polymorphisms

Single nucleotide polymorphisms (SNPs) are mutations involving a single DNA base substitution observed with a frequency of 1% or more in a given population [26, 27]. SNPs account for 90% of the inter-individual variability, and occurs approximately every 100-300 basepair in the general human population [26, 28]. This frequency makes for 100 000 amino acid differences in humans and most SNPs are inherited from one generation to the next [26]. The fact that SNPs may alter the genetic sequence in an inherited gene makes it a useful and accessible tool to perform family-based association studies [26].

1.7 Association studies

Association studies are a good way to establish a correlation between a region of the genome, SNPs and a disease [26]. By knowing the genomic position of a SNP and identifying it in a population one is able to determine an association to a phenotypic characteristic [27].

The frequency of a disease trait (i.e. SNP) is determined by comparing two groups that differ for the presence of the phenotype [26]. One of the groups will have a disease trait and the other group will either be known to be unaffected or chosen randomly from the population [27].

Association studies may be performed with a carefully pre-selected group of patients harbouring candidate genes previously linked to the disease [27]. By using a population known to have a genomic linkage to the disease one may be able to focus on biologically defined candidate genes, genes suggested by differential display experiments, or positional candidates from prior linkage investigations [27].

1.7.1 Modifier genes

Modifier genes change the phenotypic expression caused by a mutation in another gene [15]. Tumour modifier genes can be important in cancer development by either enhancing or suppressing the initiation, growth and/or progression of tumour cells [15].

There are many factors influencing genes and the development of colorectal cancer, making it difficult to study cancer development in humans due to environmental factors [29]. When studying purely the way genes influence the development of disease, animal models might be more suitable considering there are no unforeseen environmental factors affecting the results [29]. The study of Multiple Intestinal Neoplasia (min) mice is a good example of the way animal models can be used to study cancer development and modifier genes [15]. The mice harbour a germline APC mutation and have been used to study the mutation in the modifier gene ATP synthase (ATP5a1) [15].

1.7.2 The ATP synthase gene

A recent study performed at the Thomas Jefferson University showed some interesting results concerning the modifier gene called ATP synthase (ATP5a1) [15] The study

revealed that the ATP5a1 gene suppresses colorectal cancer in mice when it is mutated [15]. Mice harbouring one copy of the mutated gene had a reduction in small intestinal and colon polyps by approximately 90% [15]. The Min mice used in the study carry germline mutations in the APC gene and this is regarded as the murine version of the human FAP condition [15, 29]. The results from the study implied that the ATP5a1 gene only suppresses the tumour cells when it is present on the same chromosome as the variant APC gene (in mice they are both present on chromosome 18) [15]. However, ATP5a1 is present on chromosome 18 in humans, APC is present on chromosome 5q21-22 [15, 16].

The ATP5a1 gene encodes the α-subunit of ATP synthase, a multi-subunit enzyme which resides on the mitochondrial membrane [29]. If the ATP5a1 gene is absent this will lead to a non-functioning α synthase subunit, thus leading to apoptosis of the cell [29].

1.7.3 The candidate SNPs

The pathways involved in the control of the genome are complex and not well understood, especially considering the difficulties in disease loci identification. One way to map out the genes involved in disease is the candidate gene approach [30]. The candidate gene approach generally requires a large number of patients, which essentially have nothing in common except that they have the disease, and controls to create enough statistical power to identify an association with disease [30]. This approach is best suited for detecting genes with a minor influence on disease risk and an understanding of the mechanism underlying the disease studied are necessary [30].
Recent studies have revealed that the SNPs rs16892766, rs3802842, rs10318, rs4939827, rs4464148, rs6983267, rs7014346, rs4779584 and rs10795668 are related to an increased risk of CRC [11-13].

A study of SNP rs16892766 located on chromosome 8q23.3 of the gene EIF3H, SNP rs10795668 located on chromosome 10p14 of an unidentified gene, SNP rs6983267

located on chromosome 8q24.21 of an unidentified gene and SNP rs4939827 located on chromosome 18q21.1 of the gene SMAD7 found that there is a correlation between the SNPs and the risk of developing CRC [13].

Another study concerning SNP rs4779584 located on chromosome 15 located in between SCG5 and GREM1, and SNP rs10318 located on chromosome 15 near the gene GREM1 were significantly associated with CRC [11]. GREM1 encodes a secreted bone morphogenetic protein (BMP) antagonist and it is possible that the gene may increase tumour proliferation [11].

Finally, a study of SNP rs7014346 located on chromosome 8q24 near the gene POU5F1P1, SNP rs3802842 located on chromosome 11 near the gene LOC120376:FLJ45803 and SNP rs4939827 located on chromosome 18q21 of the gene SMAD7 were found to be correlated with risk of developing cancer [12].

2 – AIM OF THE STUDY

The aim of this study is to examine 199 Australian APC mutation positive FAP patients with a molecular diagnosis of FAP for polymorphisms in five loci on chromosomes 10p14, 8q23.3, 8q24, 11q23, 18q21 and the gene ATP5a1. This will involve assaying 9 polymorphisms in the loci described above and 18 polymorphisms in the ATP5a1 gene and statistically evaluating the effect of the SNPs and disease risk.

3 – MATERIALS AND METHODS

3.1 Materials

3.1.1 Study population

This study consisted of 199 FAP patients with germline APC mutations confirmed through molecular diagnostics of blood-extracted DNA. Out of the 199 patients, 148 was affected with polyps of which 80 developed CRC (see table 3.1). The data used in this study was collected from the Hunter Area Pathology Services (HAPS), John Hunter Hospital, Newcastle, New South Wales, Australia between the years 1997 and 2008.

A database was compiled with the following patient details; laboratory number, date of birth, gender, gene, exon, nucleotide change, consequence of mutation, polyps, age of diagnosis of polyposis/age unaffected, colorectal cancer, age of diagnosis of cancer/age unaffected, other symptoms, family cancer and consent. The form used to collect the patient details is shown in Appendix 1.

All the patients participating in the study had given informed written consent prior to the study. Ethics approval was obtained from the Human Research Ethics Committee, University of Newcastle and the Hunter New England, Newcastle, New South Wales, Australia.

Table 3.1 The samples divided into groups according to disease expression.

Group	Sample size (n)
Subject group (all samples)	199
Affected with polyps	148
Affected with CRC	80
Unaffected with polyps	51
Unaffected with CRC	119

3.1.2 Reagents, materials and equipment

The reagents, materials and equipment used in the study are listed in table 3.2-3.5.

Table 3.2 A list of the reagents used in the study.

Reagents	Concentration	Supplier
TaqMan® Universal PCR Master Mix	2 x	Applied Biosystems*
TaqMan®, SNP Genotyping Assays, Pre-Designed, SMALL-Scale	40 x	Applied Biosystems*

*Applied Biosystems, Foster City, CA, USA.

Table 3.3 A list of the materials used in the study.

Materials	Supplier
96-well PCR Microplate	Axygen® Scientific*
96-well Masterblock® plate	Greiner Bio-One **

* Axygen® Scientific, Union City, CA, USA.

** Greiner Bio-One, Frickenhausen, Germany.

Table 3.4 A list of the equipment used in the study.

Equipment	Supplier
7500 Real-Time PCR System	Applied Biosystems*
GeneAmp® PCR System 9700	Applied Biosystems*

* Applied Biosystems, Foster City, CA, USA.

Table 3.5 A list of the reagents used in the PCR setup.

Reagents	Volume in 1 reaction (μL)
TaqMan® Universal PCR Master Mix	2.5
TaqMan®, SNP Genotyping Assays, Pre-Designed, SMALL-Scale	0.125
dH_2O	1.375
DNA (50 ηg/μL)	1
Total volume into each well	5

3.1.3 DNA isolation

The genomic DNA used in the study was obtained from Molecular Genetics, Hunter Area Pathology Services, John Hunter Hospital, Newcastle, New South Wales, Australia. The DNA had previously been extracted from 10 ml EDTA blood using the "salting-out" method [31].

3.2 Methods

This study is an association study which aims to identify genetic polymorphisms that can be used to predict disease expression in individuals who are already known to be at risk of developing colorectal cancer. This risk is due a highly penetrant mutation in a gene known to be associated with colorectal cancer risk.

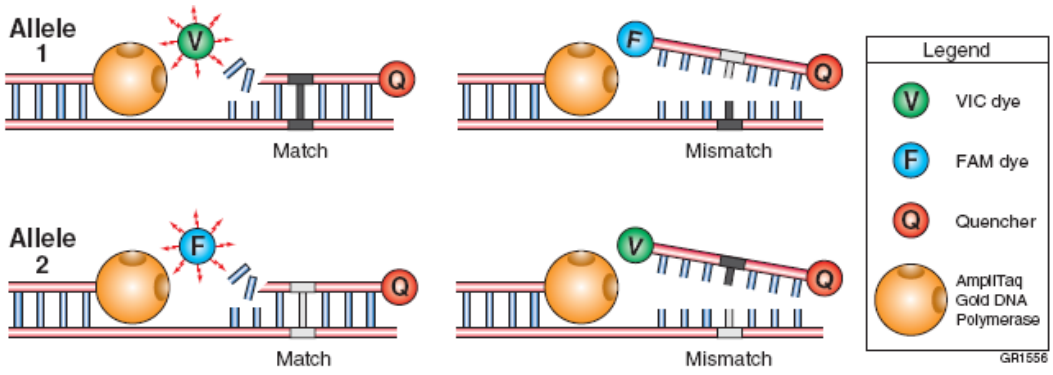
The genotypes were determined for each individual by finding the polymorphic combination for the 9 candidate SNPs in the five loci previously described. Three variable genotypes for every polymorphism were possible; homozygote wildtype (AA), heterozygote variant (Aa) or homozygote variant (aa). When a genotype was determined for a particular DNA sample, the nucleotide present at the polymorphism site was reported, i.e. in the gene SMAD7 (polymorphism rs4939827) was expressed as CC for the homozygote wildtype, CT for heterozygote and TT for homozygote variant.

The genotype frequency distribution for the 18 SNPs in the gene ATP5a1 was performed at the Australian Genome Research facility laboratory however due to insufficient time and cost effectiveness, but the results did not arrive before this thesis was due to be submitted.

3.2.1 The principle behind Real-time PCR system

The genomic DNA in this study was analysed using TaqMan ® SNP Genotyping Assays on a 7500 Real-Time PCR System and a GeneAmp® PCR 9700 System, all from Applied Biosystems.

Real-time PCR follows the general principle of the polymerase chain reaction. The difference from the traditional PCR system is the ability to monitor the PCR process as it occurs. DNA is quantified in *real-time* after each amplification cycle and the data is collected throughout the process rather than at the end of the PCR [32] .
A PCR reaction mix is used in the process. The assay contains primers to amplify a specific nucleic acid sequence and a reagent to detect and quantify the amplified target [33]. In this study the real-time PCR system is used to detect certain SNPs in a patients DNA. The SNPs are detected by using fluorogenic 5'nuclease assay (see figure 3.1). A 5' reporter dye and a 3' quencher dye are covalently linked to the probes. During the PCR extension phase, the 5' reporter dye is cleaved by the nuclease activity of the TaqMan polymerase which leads to an increase in the characteristic fluorescence of the reporter dye (see figure 3.1) By quantifying and comparing the fluorescence signals it is possible to determine the allelic content of each sample on the plate (see figure 3.2) [32]. The fluorescence is induced by light distributed from an argonion laser excitation source to all sample wells. The fluorescence is then directed to a charge couplet device camera which detects the amount of fluorescence. With every cycle the fluorescence intensity will increase in proportion with the amplification product increasing [32].



A substantial increase in...	Indicates...
VIC dye fluorescence only	homozygosity for Allele 1.
FAM dye fluorescence only	homozygosity for Allele 2.
both fluorescent signals	heterozygosity.

Figure 3.1 The figure illustrates how the probe is required to match the target sequence in order to become fluorescent. If the SNP is present this causes a mismatch in one of the probes and the sample does not become fluorescent (Figure reproduced with permission of Applied Biosystems (see appendix II) [33]).

3.2.2 Molecular analysis

All the SNPs used in this study were purchased from Applied Biosystems, Foster City, CA, USA (see table 3.2.1 and 3.2.2). Before analysing the DNA, a preliminary test was conducted to confirm the accuracy of the assays and if the reaction conditions for the real-time PCR were optimal. After a successful preliminary test, sample analysis was carried out on a 96-well optical reaction plate (Axygen Scientific). Each reaction contained 2.5 µL TaqMan® Universal PCR Master Mix, 0.125 µL TaqMan® SNP Genotyping Assay, 1.375 µL distilled water and 1 µL DNA (50 ηg/µL), with a final reaction volume of 5 µL. For every SNP, a positive control for wild type, heterozygote and/or variant genotype was provided. The plate also contained at least two negative water controls without any DNA.

All the wells were assigned to individual samples making it possible to identify the DNA samples as one of the different genotypes after the PCR and allelic discrimination had been performed. The PCR reaction was conducted on an ABI PRISM 7500 Real-Time PCR system or a GeneAmp® 9700 PCR System from Applied Biosystems. PCRs performed on the GeneAmp® PCR System 9700 had to be followed by an allelic discrimination on the 7500 Real-Time PCR system. The following thermal cycling condition was used: two initial holds at 50 °C for 2 minutes and 95 °C for 10 minutes followed by 60 cycles of 92 °C for 15 seconds and 60 °C for 1 minute.

After each PCR, an allelic discrimination was performed (see figure 3.2). If there were uncertainties about which genotype group a sample DNA belonged in due to inefficient amount of fluorescence, the sample would have to be re-analysed. If the sample persisted to give a poor amplified signal, the sample was taken out of the study for the SNP in question. Therefore a difference in sample number was obtained for each group of SNPs.

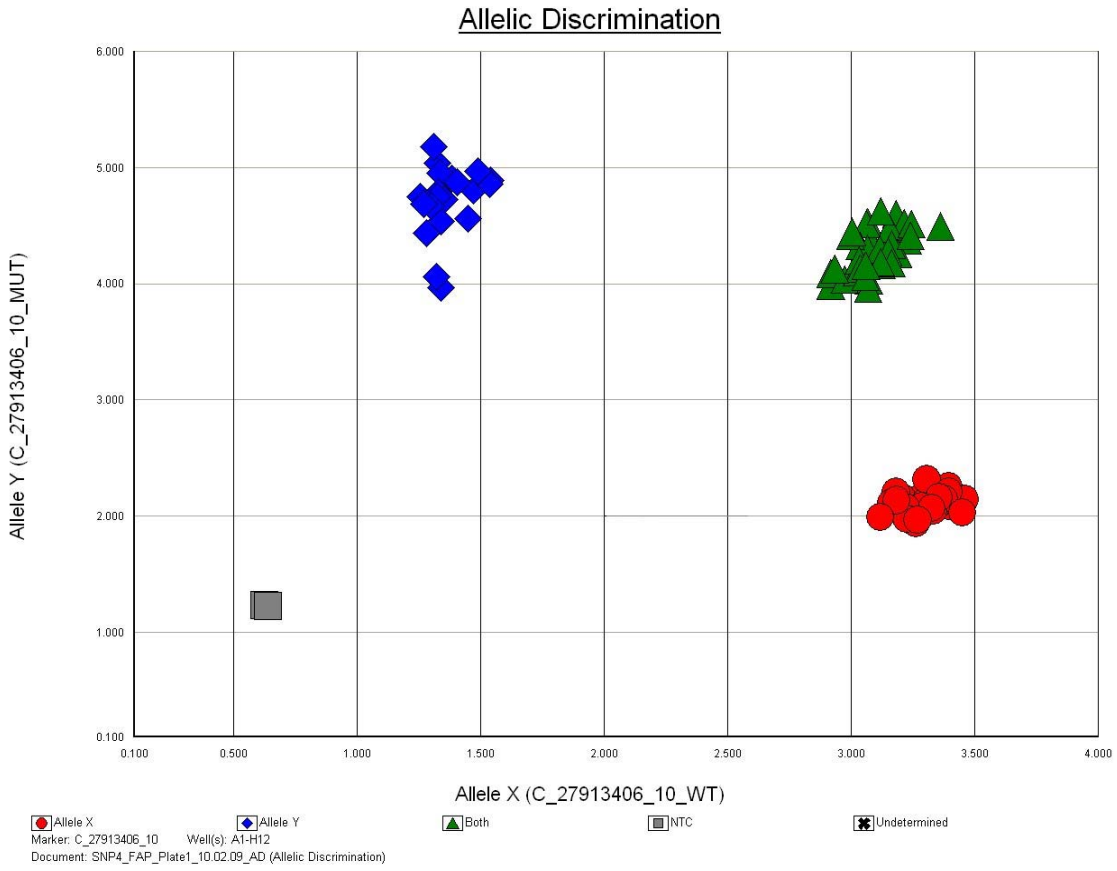


Figure 3.2 This picture shows the results from an allelic discrimination displayed in a graph. The software automatically group samples in clusters according to genotype. The diamonds represent a variant genotype sample, the triangles represent a heterozygote genotype sample and the circles represent a wildtype genotype sample. The two squares further down in the picture represent the negative water controls.

Table 3.2.1 TaqMan®, SNP Genotyping Assays, Pre-Designed, SMALL-Scale from Applied Biosystems*.

Gene**	Chromosome	Assay ID	Polymorphism (SNP)	Genotype
EIF3H	8	C_32670283_10	rs16892766	A>C
LOC120376;FLJ45803***	11	C_27503482_10	rs3802842	A>C
GREM1	15	C_12070332_20	rs10318	C>T
SMAD7	18	C_27913406_10	rs4939827	C>T
SMAD7	18	C_27989234_10	rs4464148	T>C
	18	C_29086771_10	rs6983267	T>G
POU5F1P1	8	C_29086780_10	rs7014346	G>A
SCG5, GREM1	15	C_28019826_10	rs4779584	C>T
	10	C_1779559_10	rs10795668	G>A

* Applied Biosystems, Foster City, CA, USA.

** The gene the polymorphism has been linked to.

*** Locus on chromosome.

Table 3.2.2 Table showing the context sequence for the SNPs used in the study.

Polymorphism (SNP)	Context Sequence
rs16892766 Reverse	AGACGCAAACAGTTTCAAGACTATT**[A/C]**GCTGTTAAAGGTTATGCCTTATGTC
rs3802842: Forward	GCCCTTGCAGACCCATAGAAAATCT**[A/C]**TCCCAGAAATTCACCTCATTTTAGG
rs10318: Forward	AAGATATTTGTGGTCTTGATCATAC**[C/T]**TATTAAAATAATGCCAAACACCAAA
rs4939827: Forward	TCACAGCCTCATCCAAAAGAGGAAA**[C/T]**AGGACCCCAGAGCTCCCTCAGACTC
rs4464148: Reverse	GGGGGAACAGACAGAGAAGGATGAA**[C/T]**GTGAAAAGGAAACACCCTGGTAACT
rs6983267: Forward	GTCCTTTGAGCTCAGCAGATGAAAG**[G/T]**CACTGAGAAAAGTACAAAGAATTTT
rs7014346: Forward	TCAAGATGGCTTCTGGAGTGCTACC**[A/G]**TTACATCCATGTTGTAGGCTAGAAG
rs4779584: Reverse	AGAACTTGTTGATAAGCCATTCTTC**[C/T]**GAACAGAAACCATAACTATACACAC
rs10795668: Reverse	AGAAAGAGAAAAAGTTAGATTCTTA**[A/G]**ATTCCATGATTTTATATTTCCCACC

3.2.3 Statistical Analysis

The statistical analysis was performed on the software package SPSS Version 12.0. Statistical analysis was conducted to establish whether certain SNPs segregate alone or together, with specific disease expression in FAP patients with an APC mutation. Person's chi-square (χ^2) test was performed to assess genotype frequency distribution for the different subject groups. If more than 1/5 of the groups to be compared contained less than 5 patients, a Fisher's exact test was performed. Significance of the results were accepted if $p<0.05$. Odds ratios (OR) and condifidence intervals (CI) were calculated for the groups with significant results.

Kaplan-Meier curves were used to establish differences in age of diagnosis/age unaffected with polyps and CRC within the genotype groups for the different candidate SNPs. In this study Log rank, Breslow and Tarone-Ware were tested for every Kaplan-Meier curve. The log-rank test, which emphasizes observations from later ages, the Breslow test, which gives more weight to early ages and the Tarone-Ware test that is intermediate of the two other tests, was used to examine the homogeneity of the Kaplan-Meier survival curves. The Kaplan-Meier curves that are significantly different will be the only results shown in the results section.

4 - RESULTS

4.1 Genotype frequency distribution in APC mutation positive FAP patients

The genotype frequency distribution for the different SNPs in this study was determined to establish a correlation with disease development in FAP patients. All the patients in the study had a confirmed germline APC mutation. For analysis the total population of patients were sub-divided into groups according to the following:

- Female vs. male patients.
- Affected with polyps vs. unaffected with polyps.
- Affected with CRC vs. unaffected with CRC.

In addition to wildtype, heterozygote and variant genotype, a combination of the genotypes were tested for genotype frequency distribution. The wildtype genotype was tested in a combination with heterozygote genotype against variant genotype. In addition wildtype genotype was tested against a combination of heterozygote genotype and variant genotype. This was done to establish whether the polymorphism was either a dominant or recessive modifier of disease risk.

Table 4.1 The table shows the different SNPs and states whether a significant difference in the sub-divided groups and in the Kaplan-Meier graphs was identified. It does, however, not state the genotypes for the significant differences found (for this information go to table 4.2-4.10 and figure 4.1-4.3).

SNP	Gender	Polyps	CRC	Age of diagnosis/age unaffected with polyps	Age of diagnosis/age unaffected with CRC
rs10318	No	**Yes**	No	No	No
rs16892766	No	No	No	No	No
rs3802842	No	No	No	**Yes**	No
rs4939827	No	No	No	No	No
rs4464148	No	No	No	No	No
rs6983267	No	No	No	No	No
rs7014346	No	No	No	No	No
rs4779584	No	No	No	**Yes**	**Yes**
rs10795668	No	No	No	No	No

4.1.1 SNP rs10318

For the rs10318 SNP a significant difference was observed when comparing the three different genotypes for the patients affected with or unaffected with polyps (p=0.012). In addition, another significant difference was found when comparing the wildtype genotype with the combination of heterozygote and variant genotype (CC versus CT+TT) in patients affected with or unaffected with polyps (p=0.005). Therefore, patients affected with heterozygote or variant genotype has an increased risk of developing polyps (OR 2.61 (1.33-5.15), p=0.006) (see table 4.2).

Table 4.2 Study demographics of the SNP rs10318 in APC mutation positive FAP patients according to disease expression and gender.

rs10318	CC (%)	CT (%)	TT (%)	p-value[1]	CC+CT (%)	p-value[1]	CT+TT (%)	p-value[1]
Subject group (n=196)	102 (52)	73 (37)	21 (11)		175 (89)		94 (48)	
Female (n=105)	59 (56)	32 (31)	14 (13)	p=0.083	91 (87)	p=0.203	46 (44)	p=0.212
Male (n=91)	43 (47)	41 (45)	7 (8)		84 (92)		48 (53)	
Polyps+ (n=145)	66 (46)	60 (41)	19 (13)	**p=0.012**	126 (87)	[3]p=0.111	79 (54)	**p=0.005**
Polyps- (n=51)	35 (69)	14 (27)	2 (4)		49 (96)		16 (31)	
					[2]OR=2.61 CI=1.33-5.15			**p=0.006**
CRC+ (n=78)	36 (46)	31 (40)	11 (14)	p=0.325	67 (86)	p=0.212	42 (54)	p=0.221
CRC- (n=118)	65 (55)	43 (36)	10 (9)		108 (91)		53 (45)	

[1]Comparison of genotype frequencies using Pearson's Chi-square.
[2]OR is the relative risk of developing polyps for a patient with heterozygote or variant genotype relative to patients with the wildtype genotype.
[3]Comparison of genotype frequency using Fisher's exact test.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

4.1.2 SNP rs16892766, rs3802842, rs4939827, rs4464148, rs6983267, rs7014346, rs4779584 and rs10795668

The genotype frequency distribution for the SNPs rs16892766, rs3802842, rs4939827, rs4464148, rs6983267, rs7014346, rs4779584 and rs10795668 did not reveal any statistically significant association within the different sub-groups analysed (see table 4.3-4.10).

Table 4.3 Study demographics of the SNP rs16892766 in APC mutation positive FAP patients according to disease expression and gender.

rs16892766	AA (%)	AC (%)	CC (%)	p-value[1]	AA+AC (%)	p-value[1]	AC+CC (%)	p-value[2]
Subject group (n=196)	175 (89)	19 (10)	2 (1)		194 (99)		21 (11)	
Female (n=105)	94 (89)	10 (10)	1 (1)		104 (99)		11 (11)	
Male (n=91)	81 (89)	9 (10)	1 (1)	p=1.0	90 (99)	p=1.0	10 (11)	p=0.908
Polyps+ (n=145)	132 (91)	11 (8)	2 (1)		143 (99)		13 (9)	
Polyps- (n=51)	42 (82)	9 (18)	0 (0)	p=0.121	51 (100)	p=1.0	9 (18)	p=0.091
CRC+ (n=79)	71 (90)	7 (9)	1 (1)		78 (99)		8 (10)	
CRC- (n=117)	103 (88)	13 (11)	1 (1)	p=0.908	116 (99)	p=1.0	14 (12)	p=0.689

[1]Comparison of genotype frequency using Fisher's exact test.
[2]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.4 Study demographics of the SNP rs3802842 in APC mutation positive FAP patients according to disease expression and gender.

rs3802842	AA (%)	AC (%)	CC (%)	p-value[1]	AA+AC (%)	p-value[1]	AC+CC (%)	p-value[1]
Subject group (n=196)	85 (43)	88 (45)	23 (12)		173 (88)		111 (57)	
Female (n=105)	52 (50)	40 (38)	13 (12)		92 (88)		53 (50)	
Male (n=91)	33 (36)	48 (53)	10 (11)	p=0.111	81 (89)	p=0.793	58 (64)	p=0.062
Polyps+ (n=145)	65 (45)	64 (44)	16 (11)		129 (89)		80 (55)	
Polyps- (n=51)	20 (39)	24 (47)	7 (14)	p=0.749	44 (86)	p=0.608	31 (61)	p=0.487
CRC+ (n=78)	35 (45)	34 (44)	9 (11)		69 (89)		43 (55)	
CRC- (n=118)	50 (42)	54 (46)	14 (12)	p=0.941	104 (88)	p=0.945	68 (58)	p=0.730

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.5 Study demographics of the SNP rs4939827 in APC mutation positive FAP patients according to disease expression and gender.

rs4939827	CC (%)	CT (%)	TT (%)	p-value[1]	CC+CT (%)	p-value[1]	CT+TT (%)	p-value[1]
Subject group (n=198)	55 (28)	94 (47)	49 (25)		149 (75)		143 (72)	
Female (n=106)	30 (28)	50 (47)	26 (25)		80 (75)		76 (72)	
Male (n=92)	25 (27)	44 (48)	23 (25)	p=0.984	69 (75)	p=0.939	67 (73)	p=0.860
Polyps+ (n=147)	41 (28)	70 (48)	36 (24)		111 (76)		106 (72)	
Polyps- (n=51)	15 (29)	23 (45)	13 (26)	p=0.952	38 (75)	p=0.887	36 (71)	p=0.835
CRC+ (n=80)	23 (29)	35 (44)	22 (27)		58 (73)		57 (71)	
CRC- (n=118)	33 (28)	58 (49)	27 (23)	p=0.699	91 (77)	p=0.460	85 (72)	p=0.904

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.6 Study demographics of the SNP rs4464148 in APC mutation positive FAP patients according to disease expression and gender

rs4464148	TT (%)	TC (%)	CC (%)	p-value[1]	TT+TC (%)	p-value[1]	TC+CC (%)	p-value[1]
Subject group (n=197)	98 (500	81 (41)	18 (9)		179 (91)		99 (50)	
Female (n=105)	55 (52)	41 (39)	9 (9)	p=0.731	96 (91)	p=0.768	50 (48)	p=0.429
Male (n=92)	43 (47)	40 (43)	9 (10)		83 (90)		49 (53)	
Polyps+ (n=146)	71 (49)	63 (43)	12 (8)	p=0.541	134 (92)	p=0.449	75 (51)	p=0.596
Polyps- (n=51)	27 (53)	18 (35)	6 (12)		45 (88)		24 (47)	
CRC+ (n=79)	40 (51)	35 (44)	4 (50	p=0.254	75 (95)	p=0.104	39 (49)	p=0.839
CRC- (n=118)	58 (49)	46 (39)	14 (12)		104 (88)		60 (51)	

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.7 Study demographics of the SNP rs6983267 in APC mutation positive FAP patients according to disease expression and gender.

rs6983267	TT (%)	TG (%)	GG (%)	p-value[1]	TT+TG (%)	p-value[1]	TG+GG (%)	p-value[1]
Subject group (n=197)	55(28)	82 (42)	60 (30)		137 (70)		142 (72)	
Female (n=106)	25 (24)	45 (42)	36 (34)	p=0.285	70 (66)	p=0.249	81 (76)	p=0.143
Male (n=91)	30 (33)	37 (41)	24 (26)		67 (74)		61 (67)	
Polyps+ (n=146)	40 (27)	59 (41)	47 (32)	p=0.666	99 (68)	p=0.371	106 (73)	p=0.777
Polyps- (n=51)	15 (29)	23 (45)	13 (26)		38 (74)		36 (71)	
CRC+ (n=79)	23 (29)	33 (42)	23 (29)	p=0.895	56 (71)	p=0.738	56 (71)	p=0.661
CRC- (n=118)	31 (26)	50 (43)	37 (31)		81 (69)		87 (74)	

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.8 Study demographics of the SNP rs7014346 in APC mutation positive FAP patients according to disease expression and gender.

rs7014346	GG (%)	GA (%)	AA (%)	p-value[1]	GG+GA (%)	p-value[1]	GA+AA (%)	p-value[1]
Subject group (n=198)	79 (40)	84 (42)	35 (18)		163 (82)		119 (60)	
Female (n=106)	36 (34)	49 (46)	21 (20)	p=0.184	85 (80)	p=0.398	70 (66)	p=0.067
Male (n=92)	43 (47)	35 (38)	14 (15)		78 (85)		49 (53)	
Polyps+ (n=147)	55 (38)	65 (44)	27 (18)	p=0.480	120 (82)	p=0.665	92 (62)	p=0.226
Polyps- (n=51)	24 (47)	19 (37)	8 (16)		43 (84)		27 (53)	
CRC+ (n=80)	OR= 32 (40)	36 (45)	12 (15)	p=0.686	68 (85)	p=0.416	48 (60)	p=0.981
CRC- (n=118)	47 (40)	48 (41)	23 (19)		95 (81)		71 (60)	

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.9 Study demographics of the SNP rs4779584 in APC mutation positive FAP patients according to disease expression and gender

rs4779584	CC (%)	CT (%)	TT (%)	p-value[1]	CC+CT (%)	p-value[1]	CT+TT (%)	p-value[1]
Subject group (n=198)	127 (64)	53 (27)	18 (9)		180 (91)		71 (36)	
Female (n=106)	67 (63)	27 (26)	12 (11)	p=0.491	94 (83)	p=0.241	39 (37)	p=0.769
Male (n=92)	60 (65)	26 (28)	6 (7)		86 (93)		32 (35)	
Polyps+ (n=147)	93 (63)	40 (27)	14 (10)	p=0.937	133 (90)	p=0.719	54 (37)	p=0.854
Polyps- (n=51)	33 (65)	14 (27)	4 (8)		47 (92)		18 (35)	
CRC+ (n=80)	53 (66)	19 (24)	8 (10)	p=0.645	72 (90)	p=0.714	27 (34)	p=0.529
CRC- (n=118)	73 (62)	35 (30)	10 (8)		108 (92)		45 (38)	

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

Table 4.10 Study demographics of the SNP rs10795668 in APC mutation positive FAP patients according to disease expression and gender.

rs10795668	GG (%)	GA (%)	AA (%)	p-value[1]	GG+GA (%)	p-value[1]	GA+AA (%)	p-value[1]
Subject group (n=197)	86 (44)	83 (42)	28 (14)		169 (86)		111 (56)	
Female (n=106)	46 (43)	47 (45)	13 (12)	p=0.643	93 (88)	p=0.398	60 (57)	p=0.937
Male (n=91)	40 (44)	36 (40)	15 (16)		76 (84)		51 (56)	
Polyps+ (n=146)	61 (42)	64 (44)	21 (14)	p=0.560	125 (86)	p=0.821	85 (58)	p=0.370
Polyps- (n=51)	25 (49)	18 (35)	8 (16)		43 (84)		26 (51)	
CRC+ (n=79)	29 (37)	38 (48)	12 (15)	p=0.301	67 (85)	p=0.879	50 (63)	p=0.135
CRC- (n=118)	56 (47)	45 (39)	17 (14)		101 (86)		62 (53)	

[1]Comparison of genotype frequencies using Pearson's Chi-square.
Note: affected with polyps (polyps+), unaffected with polyps (polyps-), affected with colorectal cancer (CRC+) and unaffected with colorectal cancer (CRC-).

4.2 Kaplan-Meier survival analysis

Kaplan-Meier survival analysis showed that there was a significant difference between the age of diagnosis and genotypes for SNP rs3802842 and rs4779584. For SNP rs3802842 there was a significant difference between the age of diagnosis of polyps and the combination of the wildtype and heterozygote genotypes compared with the variant genotype. For SNP rs4779584 there was a significant difference between the age of diagnosis for both polyps and CRC with the combination of the wildtype and heterozygote genotypes compared with variant genotype.

4.2.1 SNP rs3802842

When studying the Kaplan-Meier graph for SNP rs3802842 patients with the wild type and heterozygous genotypes combined compared to the variant genotype had a significant difference in the age of developing polyps (see figure 4.1). The group with wildtype and heterozygote genotype contained 173 patients and the group with mutant

genotype contained 23 patients. The median age of diagnosis of polyps (age of which 50% of the population is free of polyps) was 39 years for the combination of wildtype and heterozygote genotype and 57 years for the mutant genotype (table 4.10).

Table 4.10 Median age of diagnosis of polyps (age of which 50% of the population is free of polyps) in APC positive FAP participants for the SNP rs3802842.

Genotype	Age of subject group[1]
Wildtype and heterozygote (AA+AC)	39 yrs (n=173)
Variant (CC)	57 yrs (n=23)

[1]Subject group includes 196 samples.

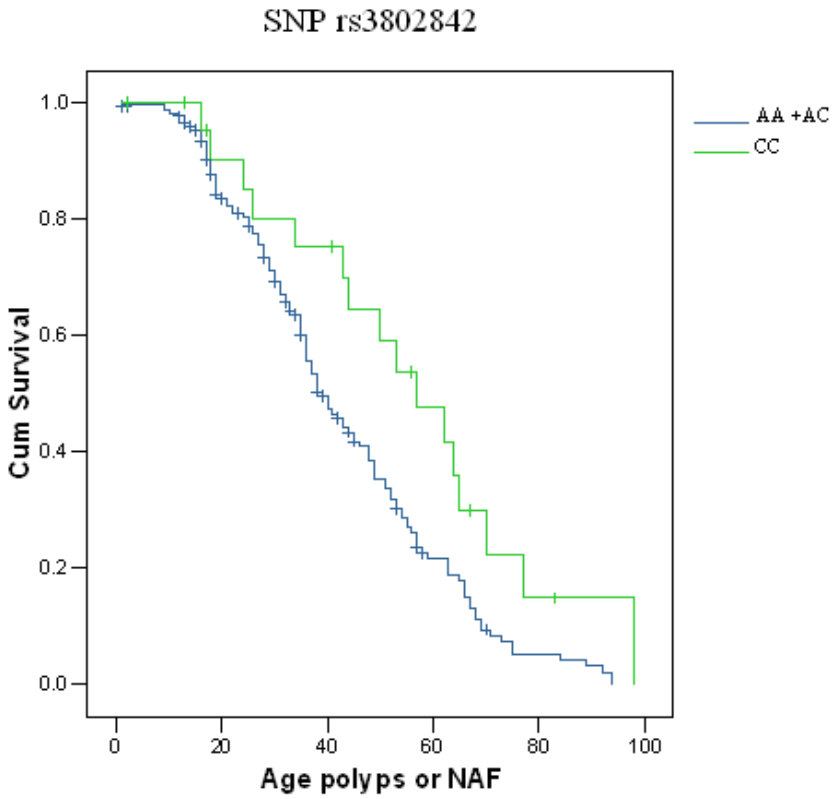


Figure 4.1 Kaplan-Meier graph for SNP rs3802842 showing the distribution of age of diagnosis/age unaffected with polyps for patients with wildtype and heterozygote genotype vs. variant genotype (log-rank test: p=0.018, Breslow: p=0.049, Tarone-Ware: p=0.029).

4.2.2 SNP rs4779584

When studying the Kaplan-Meier graph for SNP rs4779584, the combination of the wild type and heterozygous genotyped compared to the variant genotype was significantly different (see figure 4.2). The group with wildtype and heterozygote genotype contained 180 patients and the group with mutant genotype contained 18 patients. The median age of diagnosis of polyps was 43 years for the combination of wildtype and heterozygote genotypes and 28 years for the mutant genotype (table 4.11).

Table 4.11 Median age of diagnosis of polyps (age of which 50% of the population is free of polyps) in APC positive FAP participants for the SNP rs4779584.

Genotype	Age of subject group[1]
Wildtype and heterozygote (AA+AC)	43 yrs (n=180)
Variant (CC)	28 yrs (n=18)

[1]Subject group includes 198 samples.

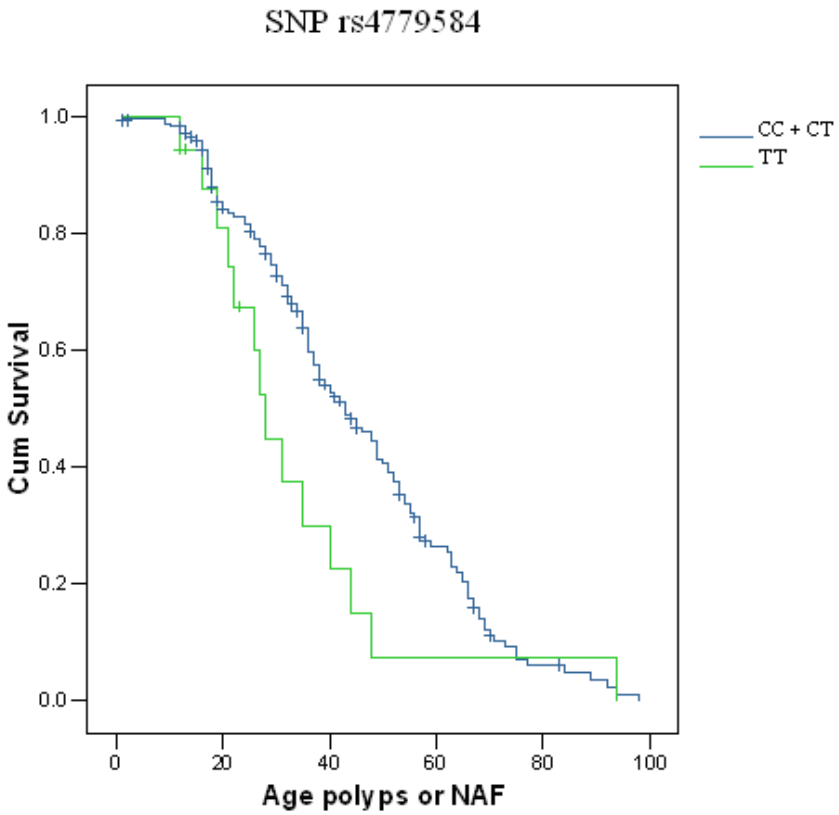


Figure 4.2 Kaplan-Meier graph for SNP rs4779584 showing the distribution of age of diagnosis/age unaffected with polyps for patients with wildtype and heterozygote genotype vs. variant genotype (log-rank: 0.062, Breslow: 0.013, Tarone-Ware: 0.014).

When studying the Kaplan-Meier graph for SNP rs4779584, the combination of the wildtype and heterozygous genotypes compared to the variant genotype was significantly different for the age of CRC development (see figure 4.3). The group with the wildtype and heterozygote genotype contained 180 patients and the group with mutant genotype contained 18 patients. The median age of diagnosis of polyps was 59 years for the combination of the wildtype and heterozygote genotypes and 44 years for the mutant genotype (table 4.12).

Table 4.12 Median age of diagnosis of CRC (age of which 50% of the population is free of CRC) in APC positive FAP participants for the SNP rs4779584.

Genotype	Age of subject group[1]
Wildtype and heterozygote (CC+CT)	59 yrs (n=180)
Variant (TT)	44 yrs (n=18)

[1]Subject group includes 198 samples

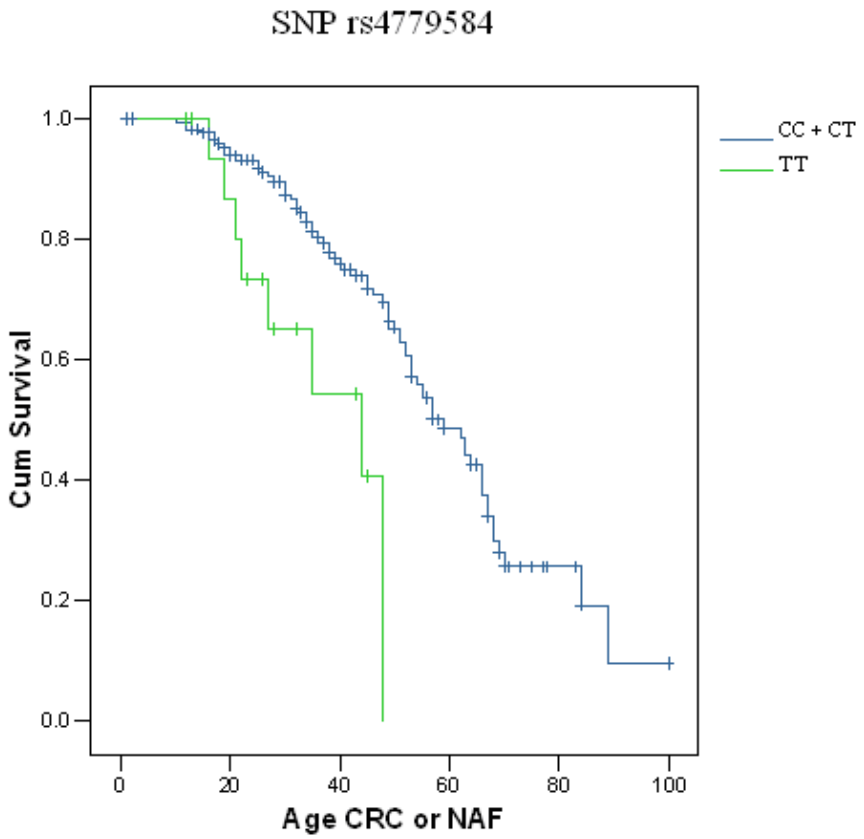


Figure 4.3 Kaplan-Meier graph for SNP rs4779584 showing the distribution of age of diagnosis/age unaffected with CRC for patients with wildtype and heterozygote genotype vs. variant genotype (log-rank: 0.002, Breslow: 0.007, Tarone-Ware: 0.004)

4.2.3 SNP rs16892766, rs4939827, rs4464148, rs6983267, rs7014346, rs10795668 and rs10318

No difference in genotype frequency was observed when the age of diagnosis of polyps or CRC was examined for SNP rs16892766, rs4939827, rs4464148, rs6983267, rs7014346, rs10795668 and rs10318.

5 – DISCUSSION

Genetic variation is the foundation for the enormous diversity in the human phenotype, and consequently the huge range of susceptibilities to common diseases, especially cancer [34]. It is generally accepted that genomic instability is the hallmark of malignancy. Cancer is a disease of abnormal gene expression due to inactivation of negative mediators and activation of positive mediators of cell proliferation [35]. By studying genetic variation it is possible to demonstrate which changes are likely to result in disease [36]. When occurring within an entire gene that contains coding and non-coding regions, single nucleotide polymorphisms (SNPs) have the potential to change protein function or its expression thereby directly contributing to disease variability observed in the expressed phenotype [37]. Association studies with polymorphisms in candidate genes that are likely to affect tumour development and progression are excellent for the purpose of identifying common genetic variants that lead to modest differences in disease risk [38]. It has been well established that the use of well defined study populations when using the candidate gene approach to search for genetic modifiers is an important testable approach that it is more likely to reveal true genetic associations between varitations in the genome and disease [39]. Genetic variation studies may not result in a clear understanding about the causative role of the identified genes, and are not always replicated in other studies [40]. This is most likely due to the fact that each genetic variant linked to complex diseases has only a modest effect on disease risk and does not adequately take into consideration gene-gene interactions yet alone gene-environmental factors that together contribute to disease risk [41].

Familial adenomatous polyposis (FAP) affects 1 in 8000 and the condition accounts for approximately 1% of all CRC [5]. Most research performed on FAP up to this point has focused on the genes, APC and MUTYH, to determine their role in disease expression in patients with this disorder. As cancer development is known to be influenced by many genetic factors, it is important to explore the possibility of different genes being involved in the the expression of disease since persons harbouring identitcal mutations in APC or MUTYH can have very different disease profiles that can not be explained by

environmental factors alone. Modifier genes have been proposed to be involved in influencing disease expression in FAP patients [42-44]. A number of recent publications that have examined the whole human genome for SNPs involved in sporadic CRC and have identified a number of SNPs that are over-represented in the CRC population in comparison to a healthy control population. The functional significance of these SNPs are not known at present but they are markers that may be important in hereditary colorectal cancer predispositions, including FAP. Therefore, the aim of this study was to determine whether these SNPs influence the age of diagnosis of polyposis and/or CRC and the risk of developing polyps and CRC among FAP patients that harboured a germline APC mutation.

A study by Jaeger *et al*. established an association between the SNP rs10318 and the risk of developing CRC [11]. Their study showed that the variant allele increased the risk of developing CRC. In concordance with this study, the results presented in the FAP population examined in this study indicate that there is an increased risk of developing polyps (OR 2.58, CI 1.33-5.15) if a patient harbours the heterozygote or variant genotype of SNP rs10318 compared to patients with the wildtype genotype. However, our results did not reveal an association for an altered risk of developing CRC and there was no significant difference observed for rs10318 and the age of diagnosis of polyps or CRC. The reason for this is likely to be due to FAP patients developing a pre-malignant phenotype relatively early in their lives, therefore these patients are a special group to evaluate differences in genotype frequency and CRC risk. The other more commonly inherited predisposition to CRC, hereditary non polyposis colorectal cancer (HNPCC) differs from FAP in that patients with HNPCC are identified when they present with disease whereas FAP patients initially present with polyps that are removed before the development of CRC. Consequently, the risk of developing CRC in a FAP population is considerably different from the most common genetic predisposition to CRC, HNPCC. For this significant reason, it is much more difficult to study CRC risk and differences in the phenotypic expression in FAP patients as many of them will have interventional surgery to reduce their risk of developing colorectal cancer. Nevertheless SNP rs10318 which is linked to the gene GREM1 is thought to increase tumour proliferation [11] and

may therefore have a relationship with colonic crypt stem cell proliferation that culminates in adenoma formation.

Results from this study have revealed that the rs3802842 SNP is protective for the early development of polyposis in FAP individuals who harbour the variant genotype compared to those carrying the wildtype or heterozygous genotypes. No significant differences were found for the rs3802842 SNP and altered risk of developing CRC or polyposis. An earlier study by Tenesa *et al*. examined the rs3802842 SNP which showed a significant difference in the risk of developing CRC in a European population [12]. However, no significant differences were observed in a Japanese population [12] although the trend of the results was similar. Differences between populations might be due to different environmental factors contributing to the differences in disease risk. Taken together, these results indicated that rs3802842 might be a marker of a population specific CRC susceptibility allele [12]. Together, our results and those of others suggest that this SNP is indeed an effect modifier as it appears to follow a similar pattern in different populations.

In addition, the results of the current study suggest that patients harbouring the variant genotype of rs4779584 develop polyps and CRC at an earlier stage compared to patients harbouring the wildtype or heterozygote genotypes. There was, however, no significant difference in genotype frequency distribution for the risk of developing polyps or CRC. A previous study performed by Jaeger *et al*. suggests that there is a correlation between the variant allele of the SNP rs4779584 and the development of CRC [11]. This study provides evidence that this SNP may be important for the early presentation of a premalignant phenotype as well as progression to disease. Identification of individuals with the variant genotype is important for the implementation of screening strategies at an early age to remove premalignant polyps that are likely to cause disease development.

The current study of a series of candidate SNPs has the potential to reveal modifiers of FAP disease expression in patients with APC mutations which will provide important information for patient management by determining individual risk of disease

development and to possibly implement personalised treatment, especially for those with early onset disease. If modifier genes influencing disease expression in FAP are confirmed and the number extended better preventative measures could be developed and implemented, leading to better patient outcomes for FAP patients who already have an increased risk of disease as they harbour an APC mutation.

One of the limitations of the study is the size of the study population. To provide stronger statistical power to confirm the findings of this study, more patients with APC mutations are required in addition to performing this analysis in an independent cohort of FAP patients. Moreover, a larger group is also required for haplotype analysis as the combination of disease associated SNPs may have an additive effect on disease expression. Performing haplotype analysis, will improve our understanding of disease pathways to help identify patients at an altered risk of disease development. FAP patients have a pre-malignant phenotypic expression in the form of polyps which are removed if discovered. Since the polyps are removed it can be difficult to study the genotypic expression of CRC in FAP patients as many of them may never present with malignancy. This presents some challenges when studying the disease expression in FAP patients. Furthermore, additional studies are required to examine the number of polyps and possible candidate SNPs as increasing polyp number, increases disease risk.

In conclusion, the findings of this study indicate that there is an increased risk of developing polyps for the patients harbouring the heterozygote or variant genotype of the rs10318 SNP. The variant genotype of SNP rs3802842 appears to have a protective effect for early polyp development in FAP individuals with APC mutations. Patients harbouring the variant genotype of the SNP rs4779584 have an increased risk of developing polyps and CRC at an earlier stage. Finally, this study has revealed a number of SNPs that possibly influence disease expression in APC mutation positive FAP patients. These findings give further insight into the molecular mechanisms involved in FAP and these SNPs may help to identify patients at an altered risk of developing disease so that the appropriate treatment can be implemented.

6 – REFERENCES

1. Jemal, A., et al., *Cancer statistics, 2008.* CA Cancer J Clin, 2008. **58**(2): p. 71-96.
2. Davidson, N.O., *Genetic testing in colorectal cancer: who, when, how and why.* Keio J Med, 2007. **56**(1): p. 14-20.
3. Enholm, S., et al., *Proportion and phenotype of MYH-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients.* Am J Pathol, 2003. **163**(3): p. 827-32.
4. Fearnhead, N.S., M.P. Britton, and W.F. Bodmer, *The ABC of APC.* Hum Mol Genet, 2001. **10**(7): p. 721-33.
5. Andersen, P., et al., *Genetiske analyser ved familiaer adenomatous polypose.* Tidsskrift for Den norske legeforening 2001(1).
6. Schulmann, K., et al., *The patient with multiple intestinal polyps.* Best Pract Res Clin Gastroenterol, 2007. **21**(3): p. 409-26.
7. Nilbert, M., et al., *Broad phenotypic spectrum in familial adenomatous polyposis; from early onset and severe phenotypes to late onset of attenuated polyposis with the first manifestation at age 72.* BMC Med Genet, 2008. **9**: p. 101.
8. Soravia, C., et al., *Desmoid disease in patients with familial adenomatous polyposis.* Dis Colon Rectum, 2000. **43**(3): p. 363-9.
9. Church, J., et al., *A desmoid tumor-staging system separates patients with intra-abdominal, familial adenomatous polyposis-associated desmoid disease by behavior and prognosis.* Dis Colon Rectum, 2008. **51**(6): p. 897-901.
10. Narayan, S. and D. Roy, *Role of APC and DNA mismatch repair genes in the development of colorectal cancers.* Mol Cancer, 2003. **2**: p. 41.
11. Jaeger, E., et al., *Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk.* Nat Genet, 2008. **40**(1): p. 26-8.
12. Tenesa, A., et al., *Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21.* Nat Genet, 2008. **40**(5): p. 631-7.
13. Tomlinson, I.P., et al., *A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3.* Nat Genet, 2008. **40**(5): p. 623-30.
14. Kairupan, C.F., et al., *Mutation analysis of the MYH gene in an Australian series of colorectal polyposis patients with or without germline APC mutations.* Int J Cancer, 2005. **116**(1): p. 73-7.
15. Baran, A.A., et al., *The modifier of Min 2 (Mom2) locus: embryonic lethality of a mutation in the Atp5a1 gene suggests a novel mechanism of polyp suppression.* Genome Res, 2007. **17**(5): p. 566-76.
16. Laurent-Puig, P., C. Beroud, and T. Soussi, *APC gene: database of germline and somatic mutations in human tumors and cell lines.* Nucleic Acids Res, 1998. **26**(1): p. 269-70.
17. Crabtree, M., et al., *Refining the relation between 'first hits' and 'second hits' at the APC locus: the 'loose fit' model and evidence for differences in somatic mutation spectra among patients.* Oncogene, 2003. **22**(27): p. 4257-65.

18. Albuquerque, C., et al., *The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade.* Hum Mol Genet, 2002. **11**(13): p. 1549-60.
19. Bertario, L., et al., *Multiple approach to the exploration of genotype-phenotype correlations in familial adenomatous polyposis.* J Clin Oncol, 2003. **21**(9): p. 1698-707.
20. Logan, C.Y. and R. Nusse, *The Wnt signaling pathway in development and disease.* Annu Rev Cell Dev Biol, 2004. **20**: p. 781-810.
21. Wilkins, J.A. and O.J. Sansom, *C-Myc is a critical mediator of the phenotypes of Apc loss in the intestine.* Cancer Res, 2008. **68**(13): p. 4963-6.
22. Wood, R.D., M. Mitchell, and T. Lindahl, *Human DNA repair genes, 2005.* Mutat Res, 2005. **577**(1-2): p. 275-83.
23. Wood, R.D., et al., *Human DNA repair genes.* Science, 2001. **291**(5507): p. 1284-9.
24. Ataian, Y. and J.E. Krebs, *Five repair pathways in one context: chromatin modification during DNA repair.* Biochem Cell Biol, 2006. **84**(4): p. 490-504.
25. Nielsen, M., et al., *Multiplicity in polyp count and extracolonic manifestations in 40 Dutch patients with MYH associated polyposis coli (MAP).* J Med Genet, 2005. **42**(9): p. e54.
26. Riva, A. and I.S. Kohane, *SNPper: retrieval and analysis of human SNPs.* Bioinformatics, 2002. **18**(12): p. 1681-5.
27. Brookes, A.J., *The essence of SNPs.* Gene, 1999. **234**(2): p. 177-86.
28. Ke, X., M.S. Taylor, and L.R. Cardon, *Singleton SNPs in the human genome and implications for genome-wide association studies.* Eur J Hum Genet, 2008. **16**(4): p. 506-15.
29. Seth, R., et al., *The putative tumour modifier gene ATP5A1 is not mutated in Human Colorectal Cancer Cell lines but expression levels correlate with TP53 mutations and Chromosomal Instability.* J Clin Pathol, 2009.
30. Kwon, J.M. and A.M. Goate, *The candidate gene approach.* Alcohol Res Health, 2000. **24**(3): p. 164-8.
31. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells.* Nucleic Acids Res, 1988. **16**(3): p. 1215.
32. *Essentials of Real Time PCR.* [cited 01.05.2009]; Available from: http://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocuments/cms_039996.pdf.
33. *Real-Time PCR vs. Traditional PCR.* [cited 01.05.2009]; Available from: http://www.appliedbiosystems.com/support/tutorials/pdf/rtpcr_vs_tradpcr.pdf.
34. Schmith, V.D., et al., *Pharmacogenetics and disease genetics of complex diseases.* Cell Mol Life Sci, 2003. **60**(8): p. 1636-46.
35. Coleman, W.B. and G.J. Tsongalis, *Molecular mechanisms of human carcinogenesis.* EXS, 2006(96): p. 321-49.
36. Girodon, E., et al., *CFTR gene mutations in adults with disseminated bronchiectasis.* Eur J Hum Genet, 1997. **5**(3): p. 149-55.
37. Sunyaev, S., et al., *Prediction of deleterious human alleles.* Hum Mol Genet, 2001. **10**(6): p. 591-7.

38. Pharoah, P.D., et al., *Association studies for finding cancer-susceptibility genetic variants.* Nat Rev Cancer, 2004. **4**(11): p. 850-60.
39. Seddon, J.M., et al., *Association of CFH Y402H and LOC387715 A69S with progression of age-related macular degeneration.* JAMA, 2007. **297**(16): p. 1793-800.
40. Rebbeck, T.R., et al., *Genetic variation and cancer: improving the environment for publication of association studies.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(12): p. 1985-6.
41. Nowell, S.A., et al., *Association of genetic variation in tamoxifen-metabolizing enzymes with overall survival and recurrence of disease in breast cancer patients.* Breast Cancer Res Treat, 2005. **91**(3): p. 249-58.
42. Crabtree, M.D., et al., *Analysis of candidate modifier loci for the severity of colonic familial adenomatous polyposis, with evidence for the importance of the N-acetyl transferases.* Gut, 2004. **53**(2): p. 271-6.
43. Crabtree, M.D., et al., *Explaining variation in familial adenomatous polyposis: relationship between genotype and phenotype and evidence for modifier genes.* Gut, 2002. **51**(3): p. 420-3.
44. Plasilova, M., et al., *Exclusion of an extracolonic disease modifier locus on chromosome 1p33-36 in a large Swiss familial adenomatous polyposis kindred.* Eur J Hum Genet, 2004. **12**(5): p. 365-71.

7 - APPENDIXES

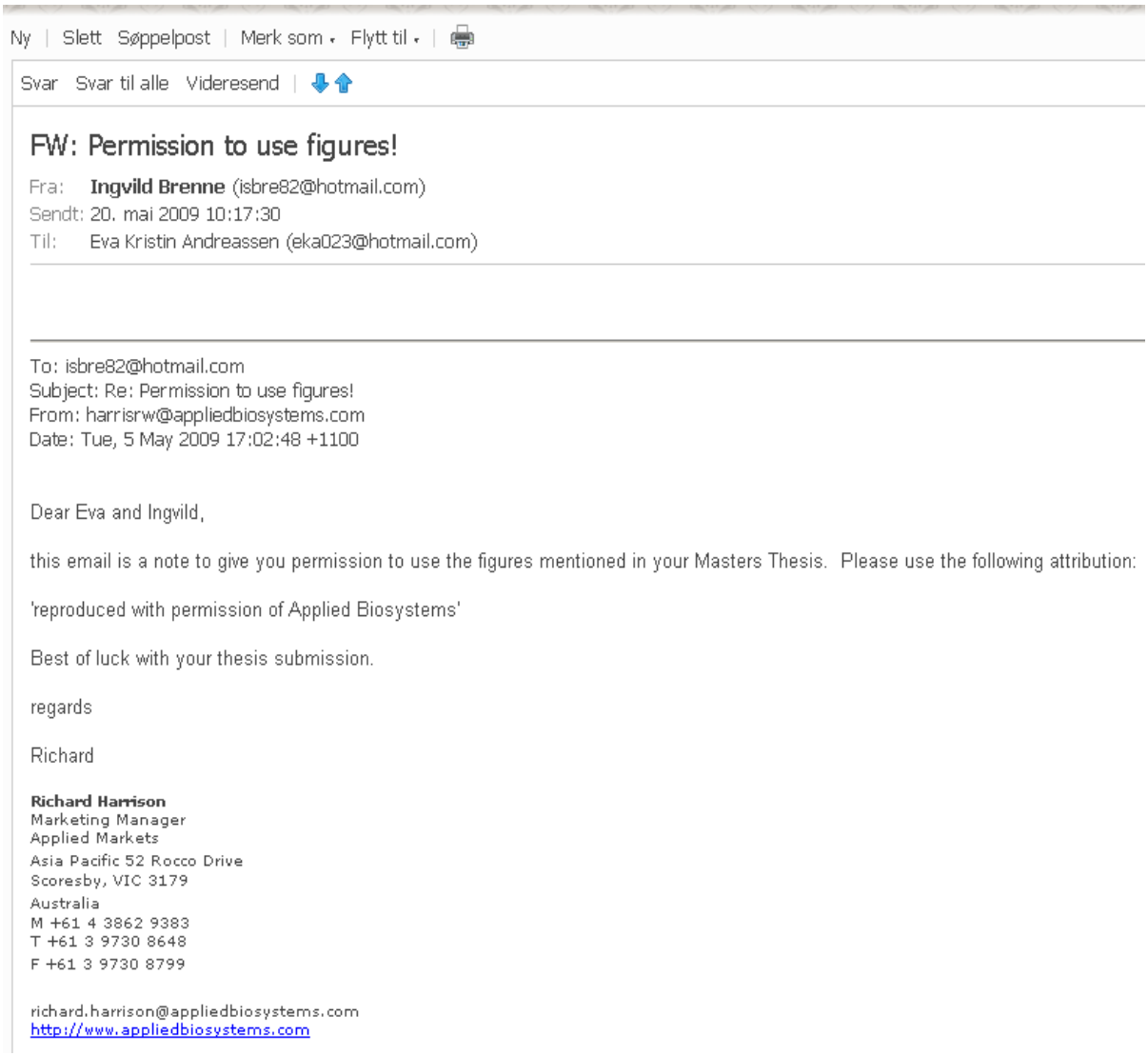
Appendix I: The form used to collect patient details.

Appendix II: Permission to use figures.

Appendix I: The form used to collect patient details.

	Lab #	Alternative lab #	DOB	Proband	Gender	Gene	Exon	Nucle.change	Cons of mutation	Polyps	Age DX/age unaffected	CRC	Age DX /age unaffected	Other symptoms	Fam cancers	Consent
1																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																
26																
27																
28																
29																
30																
31																
32																
33																
34																

Appendix II: Permission to use figures.



Ny | Slett Søppelpost | Merk som ▾ Flytt til ▾ |

Svar Svar til alle Videresend |

FW: Permission to use figures!

Fra: **Ingvild Brenne** (isbre82@hotmail.com)
Sendt: 20. mai 2009 10:17:30
Til: Eva Kristin Andreassen (eka023@hotmail.com)

To: isbre82@hotmail.com
Subject: Re: Permission to use figures!
From: harrisrw@appliedbiosystems.com
Date: Tue, 5 May 2009 17:02:48 +1100

Dear Eva and Ingvild,

this email is a note to give you permission to use the figures mentioned in your Masters Thesis. Please use the following attribution:

'reproduced with permission of Applied Biosystems'

Best of luck with your thesis submission.

regards

Richard

Richard Harrison
Marketing Manager
Applied Markets
Asia Pacific 52 Rocco Drive
Scoresby, VIC 3179
Australia
M +61 4 3862 9383
T +61 3 9730 8648
F +61 3 9730 8799

richard.harrison@appliedbiosystems.com
http://www.appliedbiosystems.com